

Search completed: July 27, 2003, 14:54:47
Job time : 4965.8 secs

QY 410 ThrLysLeuLeuTrpProLysArgLysLeuLeuHisGluGlyLeuProLysAsnHis 429
 DB 1277 GCCCTGCTGCTGTTTAAAGCGGTAAACAGCTGCTGACAGGCC-----TCC 1324
 QY 430 LysAlaAlaLysGlnAsnValArgLysGlnGlnAsnValAlaTrpLysLeuLysAla 449
 DB 1325 ACTGCTGCAGCTGCTGCTGACAGGCCGCCGCGGATGCGCGAGGCTTGAAGTCCGTTCT 1384
 QY 450 ValAspAlaPheLysSerGlyProLeuTyrgln-----ArgProGlyTyTrpSer 466
 DB 1385 GGAGACCCATTGCTTACCTAGCTGCGGGAGCCTGAACTGAGACCTGCCCATGCTCT 1444
 QY 467 AlaProGlnTrpProLeu---SerProThrProMetPhePhe----- 479
 DB 1445 GAGCCGCTGCCCCCATCCCGGGGCCAACCTCTGAACCTTATACCCAGCGTGTATCCG 1504
 QY 480 -----ProLeuGluProSerAlaProSerLysLeu 489
 DB 1505 GGGCGTGGGCTCTGCGCCGACCCGCTGCTGCTGCTGCTGCTTAT 1546
 RESULT 15
 LOCUS BC019528 1957 bp mRNA linear ROD 07-AUG-2002
 DEFINITION Mus musculus, similar to hypothetical protein FLJ20132, clone
 ACCESSION BC019528 IMAGE:4208860, mRNA, complete cds.
 VERSION BC019528.1 GI:18044530
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1957)
 REFERENCES
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.ncl.nih.gov
 CONTACT: MGC help desk
 Email: cga@ds-riemail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov
 Series: IRAC Plate: 37 Row: h Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.
 FEATURES
 Source
 1. 1957
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:28568 IMAGE:4208860"
 /tissue_type="Colon, normal, 5 month old male mouse."
 /clone_id="NCL CGAP_CO24"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 189..1586
 /codon_start=1

/product="Similar to hypothetical protein FLJ20132"
 /protein_id="AAH19528.1"
 /db_xref="GI:18044531"
 /translation="MALSAAYRFLAEBOKRYFEKLYICYDOYASLIPIYFVGFVYT
 LVHRRMNVOYLCPRLPDALMCTVAGTVGRDRGRGLYRRTLRAGLSVLLIRAST
 AVFKRPPTIDHVEYEGPRTREKRKFPENINSYKNKYVPCWPFSSILAARREGRD
 NSALKLLLELVNRSKGMLEHYMISIPLYTVQVTVIAVSYELAQIRGFLDPA
 OGKDHDTDLCPFTLLOFFEYAGMLKVAEQLINPFGDDDDFTFNFLIDRNFQVSM
 LAVDEMYPDMLMEKLDLYMDAAEAPYAFAFLQGPSFGSTFDIALAKEDQFO
 RLDDVDGPIGEVHGDFLORLPLPAGAGSVPLGRRLSLIRKRNCSVSEASTAASGAG
 AADGGVCGCGDPLDPLSLRPELEPPACRPBPAPIRPPTPEPTTVYSIIPBPAPAP
 PWLSPRIEEESPA"
 BASE COUNT 402 a 584 c 534 g 437 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 6-8e-103 Length: 1957
 Score: 1286.50 Matches: 269
 Percent Similarity: 64.79% Conservative: 64
 Best Local Similarity: 52.33% Mismatches: 142
 Query Match: 41.23% Indels: 39
 DB: 10 Gaps: 9
 US-09-622-964-3 (1-585) x BC019528 (1-1957)
 QY 1 Methrillerhrrtyrrhserglnvalaasnaalargleuglyserpheserargleu 20
 DB ATGACCGTACCTACACAGCAGAGTGGCGAATGCCCGCTTGCGGCTTCGCGAGCTG 119
 QY 21 LeuLeuCytrpArgLysSerLysLeuLeuTyrglyGlnPheLeuIlePheLeu 40
 DB CTGCTGCTGCGCGGAGAGATCTACAGCCCTGCTGCGGAGCTGTATGTTTCCG 179
 QY 41 LeuCytrpTyrrLleIleArgPheLleTyrrArgLeuAlaLeuThyGluGlnGlnLeu 60
 DB GGACCTTCACAGGACCTACAGCCGCCCTATCCCTTACTGCGCAGAAAGACAGAGCC 239
 QY 61 MetPheGluLysLeuThyLeuTyrcysaspserTyrrleGlnLeuIleProIleSerPhe 80
 DB TACTTCGAGAGCTTGTATATACGACAGCAGTCCAGCCAGCTCATCCCGCTCTTTC 299
 QY 81 ValLeuGlyPheTyrrValThrLeuValValThrArgTprPargLntTyrcLysLeu 100
 DB GTACTTGGCTCTACGACTCTGCTGTCGATCCCTGCTGAGAACCACTATGATGATG 359
 QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspLugln 120
 DB CCTCTGCGCGAGCAGCATGTCATAGTGGCTGCGACCGTGCATGGCGAGACATGCA 419
 QY 121 SerArgLeuLeuArgGlyThyLeuIleArgTyrrAlaAsnLeuGlyAsnValLeuLeu 140
 DB GGGCGCTCTACCGCGGACCGCTCATGCGACGAGGCTCTCCGCGCTGCTGATCCTT 479
 QY 141 ArgSerValSerThrAlaValTyrrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
 DB CGTCTGTTCACACAGAGCTTCAACAGGTTCCCACTATAGACACAGTGGTCAAGCT 539
 QY 161 GlyPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsnMet 180
 DB GGATTTATGACCCGAGAGAGCGCAGAGAGTTCGAGAACTTGATTCCTACACAA 599
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyLysArg 200
 DB TACTGGGTGCGCTCTGATGTTCTCCGACGCTGCGACGCGCGCGAGAGGCGCG 659
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgPheLnt 220
 DB ATCCGCGACACAGACTGCCCTTAAGCTTCTACTAGAGGCTGAATGTGTCGAGCA 719
 QY 221 CysGlyHisLeuTyrrAlaTyrrAspTrpLleSerLleProLeuValTyrrThGlnValVal 240
 DB TGTGGAGTGTCTTACTACGAGCTGATGATATACCCCTGCTCATACACAGAGTATC 779

AUTHORS
JOURNAL

Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mdickson@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LILN at: <http://image.liln.gov>
Series: IRAC Plate: 58 Row: 9 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source

Location/Qualifiers
1..1956
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:37621 IMAGE:4989959"
/tissue_type="Colon, normal, 5 month old male mouse."
/clone_lib="NCL_CGAP-Co24"
/lab_host="DH10B"
/note="Vector: PCMV-SPORE6"
176..1573
/product="Unknown (protein for MGC:37621)"
/protein_id="AA31186.1"
/db_xref="GI:21411100"
/db_xref="LocusID:212989"
/translation="MALSAYRFLLAEEOKRREKLVICDOYASLIPVSVLQSVYT
LVVHWMNOYLCMLPDALMCIVAGVHGRDGRRLYRRTLMARSLAVLIRSVT
AVKRRPPTIDHYVEAGFMRERKKRRENTSSNRYKRWCVFSSLAQARREGRLD
NSAKLLEELNVEFSKCMLEPHYDHSIPLYTOVTVTAIVSYFLACILIDNFOVSA
OGYKHRTDLCPITFLDOFFYAGLKYAQLINFGSDDDFEENFLIDNFOVSA
LAVDKATDGLAEKEDLYWDAEARAPYAAATFLQOPSGFSTFEDIALAKEDMOFO
RUDVGDPGLGEVHGDFLQRLPAGASVGPLRGLSLRRKSCVSEASTASCCGAG
AADGGVCGCDPLDPSLRPELEPRACPRPAPIRPPTPEPTTVSIPGRAPAP
PWLSPRIGEEESPAP"

CDS

BASE COUNT 411 a 577 c 530 g 438 t
ORIGIN
Alignment Scores:
Pred. No.: 6.79e-103 Length: 1956
Score: 1286.50 Matches: 269
Percent Similarity: 64.79% Conservative: 64
Best Local Similarity: 52.33% Mismatches: 142
Query Match: 41.23% Indels: 39
Gaps: 9
US-09-622-964-3 (1-585) x BC031186 (1-1956)

OY 1 MethrillerhrryrrhSerglnValaalaanaaargleuglySerPheserArgleu 20
Db 47 ATGACCGCCTACACAGCAGAGTGGCGAATCGCCCTCGGAGGCTCTCGAGCTG 106
OY 21 LeucystyrrParglySerlletyrrlyleuueuryrglyglupheleuilepheleu 40
Db 107 CTGCTGCTGCTGCGCCGAGAGCATCTACAGCTCTGTCGCGAGCTGTATGTTCTCG 166
OY 41 LeucystyrrTyrrleleargpheelletyrrargleualaaleuThrgluglIngleu 60

Db 167 GGACTCTACATGGCAGTAAGCCCGCCTATCGCTTACTGGCAGAGAGCAGAGCC 226
OY 61 MethpuglyuylleuThleuThyrcysAspserTyrrleleuileuileProliSerphe 80
Db 227 TACTTCGAGAAAGCTGTATATCTGACAGCAGTGGCAGCCAGCTCCCTCTCTTC 286
OY 81 ValleuglypethyrrValThleuValValThraagtrpTPasnglTyrcylusnleu 100
Db 287 GTACTTGGCTCTACGTACTGTGTGTGATCGCTGTGAGACCACTACTATGATG 346
OY 101 ProtrpPaoaPargleuMetSerleuValSerllyPheValgluglyLysaspIngn 120
Db 347 CCTTCCGCGAGCAGCATATGTGATGCTGCTGGCAGCTGCGATGCGGAGAGATGCA 406
OY 121 SerargleuauArgargThleuileargTyrrAlaasnleuglyAsnValleuileu 140
Db 407 GGCGCGCTACCGCGCGCGACCTATGCGTACCGAGGCTCTCCGCGCTGATCTT 466
OY 141 ArgserValSerThraAlaValTyrrLysargPheProserAlaInHlsleuValGlIna 160
Db 467 CGTTCTGTTCAGACAGCAGCTTCGAAAGGTTCCCACTATGACACACAGCTGTGAGCT 526
OY 161 GlyPheMetThraProAlaInHlsleuInleuInlyleuSerleuProHlsMet 180
Db 527 GGATTATGACCGCAGAGAGCGCAGAGCTTCGAAATTCGCTCTACACAA 586
OY 181 PheTrpValProTrpValTrpPheAlaasnleuSerMetLysAlaTrpLeuGlyArg 200
Db 587 TACTGGGTGCGCTGCGATGATGTTCTGAGCTGCGACGCGCGCGAGAGGCGCC 646
OY 201 IlargAspProleuLeuInleuInSerleuInleuInleuInleuInleuInleu 220
Db 647 ATCCGGGACAAAGTCCCTAAAGTGTCTAGAGAGCGTGAATGTTTGGAGCAA 706
OY 221 CysGlyHlsleuTyrrAlaTrpPheAlaasnleuSerleuProleuValTyrrPheVal 240
Db 707 TGTGGATGCTGTTTACAGTACAGTGTGATGATGATGATGATGATGATGATGATG 766
OY 241 ThValAlaValTyrrSerPheleuThrcysleuValglYargInPheleuAsnPro 260
Db 767 ACTATGCGAGTGTACAGTCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826
OY 261 AlaValAlaTyrrProclYHlsleuAspLeuValProValPheThrPheleuIn 280
Db 827 GCAGAGGCTACAAAGACCAACCCCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTG 886
OY 281 PhePhePheTyrrValglYrrleuValAlaInleuInleuInleuInleuInleu 300
Db 887 TCTCTCTCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
OY 301 AspAspAspPhegluThraasnTrpIleValaAspArgasnleuInValSerleuIn 320
Db 947 GACGACGAGCAGCTTGTAGACCACTTCTTATGACCGCAGCTTCAGTGCATGCA 1006
OY 321 AlaValAspGluMetHlsleuAspLeuProArgMetgluProAspMetTyrrPasnly 340
Db 1007 GCTGTAGTATGATATATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1066
OY 341 ProgluProgluProTyrrThraAlaAspLeuValaAspLeuValaAspLeu 359
Db 1067 GCAGAGCTCGCGCGCGCTACACCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTG 1126
OY 360 MetGlySerThrPheasnIleSerleuAsnlysgluMetgluPhegluProasnIn 379
Db 1127 CAGGCTCCACCTTGTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1180
OY 380 GluAspGluAspPheValaAspPheValaAspPheValaAspPheValaAspPhe 394
Db 1181 CTGACGCGCGCGGATGCGCGCGCTGAGAGAGTTCACGGGACTTTCACACGCTCTG 1240
OY 395 GlyleuInSerHlsAspHlsAspProAlaInleuInleuInleuInleuInleu 409
Db 1241 CCGCGCGCGCGCGGCTTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1276

JOURNAL Hum. Mol. Genet. 7 (9), 1517-1525 (1998).
MEDLINE 98367043
PubMed 9700209
REFERENCE
AUTHORS 2 (bases 1 to 1506)
TITLE Stohr, H., Marquardt, A., Nanda, I., Schmidt, M. and Weber, B. H.
JOURNAL Three novel human VMD2-like genes are members of the evolutionary
MEDLINE highly conserved RFP-TM family
PubMed 22027749
Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
JOURNAL 12032738
REFERENCE 3 (bases 1 to 1506)
AUTHORS Stohr, H., Marquardt, A. and Weber, B. H.
TITLE Three novel human VMD2-like protein genes are members of the
JOURNAL evolutionary highly conserved RFP family
REFERENCE 4 (bases 1 to 1506)
AUTHORS Stohr, H., Marquardt, A. and Weber, B. H.
TITLE Direct Submission
JOURNAL Submitted (30-OCT-2001) Human Genetics, University of Wuerzburg,
Blozentrum, Am Hubland, Wuerzburg 97074, Germany
FEATURES
source
1.1506
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q14.2-q15"
173.1369
/note="VMD2L3: putative membrane protein; Pfam01062;
worm_family_8"
/codon_start=1
/product="vitelliform macular dystrophy 2-like protein 3"
/protein_id="AA076997.1"
/db_xref="GI:2173484"
/translation="MTPTYSKYANATFFGRHLIKRNGSYIKLLRETFEFAVLYT
ATSIYRRLITGVQKRYEKLSTICDVAEQIPVTVLGFVTLVNVNRMNDFVLYP
PRLMFLISSVSHSGDEHGRRLRLRLRLVNLSTLIFRSVTAAYKRPMDHYVEA
GEMTDERLFLNHLKSPHLKYVPFTFENLAKRNEGRIDSDVDSGLMENRNR
SWCSLIFGDMVGIPLYVQVTVLAVYEFACILGROPLDITKYGAGSDILYPIE
TLLOFEFFAGMLKVAEOLINPRGEDDDDEFTWCIDRNOYSLAVDEHMSLPRMK
DIYMDSDAARPPYTLAADYICPLPSLGSIVONGKQKPMKEMEDIKIPLPQPOCA
KSDPGG"
BASE COUNT 416 a 341 c 330 g 419 t
ORIGIN
Alignment Scores:
Pred. No.: 7.13e-104 length: 1506
Score: 1296.00 Matches: 238
Percent Similarity: 75.853 Conservative: 51
Best Local Similarity: 62.478 Mismatches: 90
Query Match: 41.548 Indels: 2
Gaps: 1
US-09-622-964-3 (1-585) x AF440758 (1-1506)
QY 1 MethThIleThrTyThSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 173 ATGACTGACCTACTCTCCAGTAAGTACCAATGCAACATTTTGGATTTCATGAGTTA 232
QY 21 LeuLeuGlySerPheGlySerIleTyLeuLeuTyGlyGlnPheLeuIlePheLeu 40
DB 233 CTCCTCAAGTGGAGGAGCATCTACAACTACTCTGACAGGAAATTTATTTGTTTCT 292
QY 41 LeuGlyTyTyIleIleArgPheIleTyArgLeuAlaLeuThrGlnGlnGlnLeu 60
DB 293 GTTCTTTATACAGCAATTAAGTTGGTGTACAGATTGTATACAGAGGTCACAAAACGT 352
QY 61 MetPheGlnTyLeuThrLeuTyCysAspSerTyIleGlnLeuIlePheIleSerPhe 80
DB 353 TACTTTGAAAAATATATCAATTTACTGTGACAGATATGCTGAACAAATTCAGTAACCTT 412
QY 81 ValLeuGlnPheTyValThrLeuValValThrArgTyrPheAsnGlnTyGlnAsnLeu 100
DB 413 GTGCTGTGGTGTATGTACTCTGTGTAACGATGCGTGAACCAAGTTTGATGATTTG 472

QY 101 ProThrProAspArgLeuMetSerLeuValSerGlyPheValGlyGlyLysAspGlnGln 120
DB 473 CCCTGGCCACAGACAGTAAGTTCCTGATCTAGCAGTGTTCACGGAAACGACGACAC 532
QY 121 SerArgLeuLeuArgArgThrLeuIleArgTyAlaAsnLeuGlyAsnValLeuIleLeu 140
DB 533 GGGGCGCTGTTGAAGAGGAGCGTATCGCTAGCTCAATCTCACCTCGCTCATCTTT 592
QY 141 ArgSerValSerThrAlaValTyTyArgPheProSerAlaGlnIleLeuAla 160
DB 593 CGCTCGGTGGACACTGCTGTGTACAAAAGATTTCACATGACACGACCGTGTGAAGCA 652
QY 161 GlyPheMetThrProAlaGlnIleLysGlnLeuGlnLysLeuSerLeuProIleAsnMet 180
DB 653 GGTTTATGACACACAGATGAAGAAATTTATTCACCCCTCAAGTCTCTCATCTGAAA 712
QY 181 PheTyValProThrValTyPheAlaAsnLeuSerMetLysAlaThrLeuGlyLysArg 200
DB 713 TATTGGTTCATTCATCTGTTGTGAATCTTCACAACTGAACCCGGAATGAGTAGA 772
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220
DB 773 ATCAGACAGAGTGTGATCTGCATTCATGATGACATGATGATGATGATGATGATGAT 832
QY 221 CysGlyHisLeuTyAlaTyArgAspThrIleSerIleProLeuValTyThrGlnValAla 240
DB 833 TGCAGCCCTTTATTCGCTTATGACCTGGTGGGATTCCTCGTGTATACACCAAGTTGTC 892
QY 241 ThrValAlaValTySerPheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB 893 ACTCTGTGCTATACCTCTCTTCTTGGCTGCTGATGACGACGACGATTTTGATGCC 952
QY 261 AlaLysAlaTyProGlnHisGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
DB 953 ACCAAAGCGCTACGACAGGAGTACTGTGATCTTTCATCTCCATCTCCCTCTACAA 1012
QY 281 PhePhePheTyValGlyTyPheLysValAlaGlnGlnLeuIleAsnProPheGlyGln 300
DB 1013 TTCTCTCTCTATGACGAGATGGTTAAGTAGACAGACGACCTTATCAACCTTTTGAGCA 1072
QY 301 AspAspAspAspPheGlnThrAsnThrPheValAlaAspArgAsnLeuGlnValSerLeuLeu 320
DB 1073 GATGATGATGATTTGAAACTAACGTGCTGATGACAAATTTGACAGTCTCTCTTTA 1132
QY 321 AlaValAspGlnMetHisGlnAspLeuProArgMetGluProAspMetTyTyTrpAsnLys 340
DB 1133 GCTGTGACGAAATGACATGACCTTACCCCAAGATGAAGAAAGACATTTACTGGAGACAT 1192
QY 341 ProGlnProGlnProProTyTyThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
DB 1193 TCTGCTGCTGCCCCACCATACATGATGACACTGCTGATCTGACATACCTCATTTCTG 1252
QY 361 GlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGlnPheGlnProAsnGlnGln 380
DB 1253 GGGTCACAA-----GTCCAGATGGGGAACAGATGCCCTAAGAAATGAGTGAAGATGGA 1306
QY 381 Asp 381
DB 1307 GAT 1309
RESULT 14
BC031186
LOCUS BC031186 1956 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, clone MGC:37621 IMAGE:4989959, mRNA, complete cds.
ACCESSION BC031186
VERSION BC031186.1 GI:21411099
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1956)

TLQEFYAGMLKVAEQILNPFGEDEDDPEFNKICIDRNOVSLAVDENKMSLPKMK
 DIYWDLSARPPYTLAAADYCIPISEFGVOMKISSDPFDEEMIMDEKHSRSMI
 RKVRKELSHHESSPRRSYRROTSDSMIPRODLSARLOLVSNNPSPAPPTV
 KSKPEGSPPLHSMGELSTIRETSQISLSTLPQSSVRSPTKMPYEPVLTAA
 EAPVPSGCHHDSATSLUSSEFTVQOQOQPKMSILSPSEKPEPSPOT
 VASABENFINCEEDPGDFTFLKMSLPGFLGSSHTSLGNLSPDPSOALLIDETS
 SETSINIVAGSRVSDMLYLMENDTYETDIIEINKREESPK*
 BASE COUNT 689 a 632 c 563 g 616 t

Alignment Scores:

Pred. No.: 9,75e-107 Length: 2500
 Score: 1332.00 Matches: 276
 Percent Similarity: 63.52% Conservative: 67
 Best Local Similarity: 51.11% Mismatches: 159
 Query Match: 42.69% Indels: 38
 Gaps: 8

US-09-622-964-3 (1-585) x AK096459 (1-2500)

Oy 1 MethrilethiThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 |||||
 Db 229 ATGACTGCTACTCTCCAGTAAGTACCAACTTTTGGATTTCATAGGTTA 288
 Oy 21 LeuLeuGlySerPheGlySerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 |||||
 Db 289 CTCCTCAAGTGGAGGACGACATCTACCAACTCTGACAGGAAATTAATGTTTGGCT 348
 Oy 41 LeuGlySerThyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 |||||
 Db 349 GTTCTTATACAGCAATAGTTGGCTACAGATTGTTACTTACAGAGATGCAAAAACGT 408
 Oy 61 MethLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 |||||
 Db 409 TACTTTGAAAATATCAATTAATCTGACAGATATGCTGACCAATATCCAGTAACCTTT 468
 Oy 81 ValLeuGlyPheThyValThrLeuValAlaThrArgTrpAsnGlnTrpGlnAsnLeu 100
 |||||
 Db 469 GTGCTGGGTTTATGTTACTCTGCTGACAGCAACGATGGTGAAACAGTTTGTGAATTTGG 528
 Oy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLeuAspGln 120
 |||||
 Db 529 CCTGGCCACAGAGGCTTAATGTTCCCTACCTACAGAGTACAGGAAAGCCAGACGAC 588
 Oy 121 SerArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 140
 |||||
 Db 589 GGGCGCTGCTTACAGAGACGCTGATGCTGCTACCAATCTCACCTCCCTGCTCATCTTT 648
 Oy 141 ArgSerValSerThrAlaValThyTrpAspArgPheProSerAlaGlnHisLeuLeuAla 160
 |||||
 Db 649 CGCTGGTGACACGCTGCTGACAAAGATTTCCACACATAGGACACGCTGGTGAAACA 708
 Oy 161 GlyPheMetThrProAlaGlnHisGlnLeuGlyLeuLeuSerLeuProHisAsnMet 180
 |||||
 Db 709 GGTTTATGACACACAGATGAAGAAATTAATTCACACCACTCACTCACTCATCTGAAA 768
 Oy 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLeuAlaTrpLeuGlyArg 200
 |||||
 Db 769 TATGGGTTTCATCATCTGTTGGAATCTGCACTGAAGCCCGAATGAAGTGA 828
 Oy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThcGln 220
 |||||
 Db 829 ATCAAGACAGCTGGTATGCTCAATCATGATGATGAATGAATGATGATGATGATGATG 888
 Oy 221 CysGlnHisLeuTrpAlaValTrpAspTrpIleSerLeuProLeuValThyTrpGlnVal 240
 |||||
 Db 889 TGCAGCCCTTATTCGGTATGATGGGTTGGATTCGGCTGGTTTACACCCAGGTTGTC 948
 Oy 241 ThrValAlaValTrpSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 |||||
 Db 949 ACTCTCTGCTATACCTCTCTTCTTGGCTGATGATGATGATGATGATGATGATGATG 1008
 Oy 261 AlaLeuAlaValTrpProGlnHisGlnLeuAspLeuValValProValPheTrpPheLeuGln 280

Db 1009 ACCAAGGCTACGACGGGCACTGATCTTACATTCATCCATTCACCTCTACAA 1068
 |||||
 Oy 281 PhePhePheThyValGlyTrpLeuValAlaGlnGlnLeuLeuLeuLeuLeuLeuLeu 300
 |||||
 Db 1069 TTCCTCTCTTATGCGATGATGCTTAAGTACAGACGACGCTTACCAACCTTTGGACA 1128
 Oy 301 AspAspAspPheGlnThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320
 |||||
 Db 1129 GATGATGATGATTTTGAACACTTACCTGGGCTGTATACAGAAATTTGACAGTCTCTTTTA 1188
 Oy 321 AlaValAspGlnMetHisGlnAspLeuProArgMetGlnProAspMetTrpTrpAsn 340
 |||||
 Db 1189 GCTGTGACGAAATGACATAGCTTACCCCAAGATGAAGAGACATTTACTGGACGAT 1248
 Oy 341 ProGlnProGlnProProTrpThrAlaAlaSerAlaGlnPheArgAlaSerPheMet 360
 |||||
 Db 1249 TCTGCTGCTGCGCCACATACATGCTGACAGCTGCTACTACTGATACCTCATTTCTG 1308
 Oy 361 GlySerThrPheAsnIleSerLeuAsnLeuGlnLeuMet-----GlnPheGlnPro 377
 |||||
 Db 1309 GGGTCAACAGCTCCAGATGGGGCTGTGCGGCTCGGACTTCTCTGACAGAGAGTGGCTGG 1368
 Oy 378 AsnGlnGlnAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 395
 |||||
 Db 1369 GATTATGAGAGAGATGGCCATCGCATTCATGATGAAGAGATCAAGCGCTTCTGAGT 1428
 Oy 396 LeuGlnSerHisAspHisAspHisAspHisAspHisAspHisAspHisAspHisAsp 412
 |||||
 Db 1429 GCCCAGACACACCCCTCCAGCCAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 418
 |||||
 Oy 413 -----LeuTrpProValArgGlnSerLeuLeu-----His 422
 |||||
 Db 1489 TCCATGTTCTTACCCGAGATGACCTGACCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1548
 Oy 422 SGIUGLYLEUPROLYSASPHISLYSALALALYSGLINASNVALARGLYGLINUSPRA 442
 |||||
 Db 1549 ABAACCCGCCAGAGGCTCACCCACCTGGAAGA--AATCGTGTCTCCCAAGAGAGAG 1605
 Oy 442 nlysalatrpLyLeuLyAlaValAspAlaPheLySerGlyProLeuTrpGlnArgPr 462
 |||||
 Db 1606 CCGA-----CGTGCATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1653
 Oy 462 oGly-----TrpTrpSerAlaProGlnTrpProLeuSerProTrpPheMetPh 478
 |||||
 Db 1654 ACCGACAGACACTTTCACAGAGCTGACCCACAGCTCAGTGTGAGAACTTCC----- 1708
 Oy 478 ePheProLeuGlnProSerAlaProSerLySerHis--SerValThrGlyIleAspTrp 498
 |||||
 Db 1709 -----CCATCAAAATGCCACTGTGTAACCTGAGGTATGATTCACA 1746
 Oy 498 ysaAspLySerLeuLyThrValSerSerGlyAlaLyLySerPheGlnLeuLeu 516
 |||||
 Db 1747 GCAGCCGAAAGCAGCAGTCCACATCAGGGGGCTACCCATGATTCGGCTACCTC 1802
 |||||

RESULT 13
 AF440758
 LOCUS 1506 bp mRNA linear PRI 12-JUL-2002
 DEFINITION Homo sapiens vitelliform macular dystrophy 2-like protein 3 mRNA,
 complete cds.
 ACCESSION AF440758
 VERSION AF440758.1 GI:21734843
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1506)
 AUTORS Marguardt, A., Stohr, H., Passmore, L.A., Kramer, F., Rivera, A. and
 Weber, B. H.
 TITLES Mutations in a novel gene, VMD2, encoding a protein of unknown
 properties cause juvenile-onset vitelliform macular dystrophy
 (Best's disease)

Db	68379	CTACATGGTCTGTACCCAGCAGCTGAAGGTTGTTGAGGGGTGGGAGGGCTGAAAACAG	68322
QY	366	-----	366
Db	68319	AACGATAAGCATAGACCTTGTCACAGGAATGACAAATTATGAGGAGACTCAACC	68260
QY	366	-----	366
Db	68259	CAGCCTCAACCTGTGATACAGGTACAAAGTACTGATGTCAGAAAAGGACAGAAC	68200
QY	366	-----	366
Db	68199	ATGSAACACAGTCATCTTTGTCTGCTGGGAGGGCGCTTCACGTGGTCTGAGCTGAG	68140
QY	366	-----	366
Db	68139	CCATGGAACATGGGAGAGATGTGAACCTGGGCAAGGGCCACTCTCTGGTAGTA	68080
QY	366	-----	366
Db	68079	AGCTTCCCTTCGAGGGTAAAGGTCTGGGGCTCCCGGAGTCGTGTGCTAGAAATCA	68020
QY	366	-----	366
Db	68019	ATTTCCTTTGTGGATCTCACTCCCAATGGAACCAAAATCTCGCATTCAGAGATC	67960
QY	366	-----	366
Db	67959	ACTCATGGCCCTCATCTGAACCACTCATGCCAGGGCAGCAGTGTCTTGACTGCCTGAG	67900
QY	366	-----	366
Db	67899	TGAGGGGTTTACAGGGGAAAGTANGATGAGAGGGCTTTACAGCCAGCGGGGGTGG	67840
QY	366	-----	366
Db	67839	TTGCGGGGGTGGATGTTAACTCTGTGCAGAGAGGGAATCAACAACAGTGAGTGAAGCTG	67780
QY	366	-----	366
Db	67779	GGCCTGAGAGGATCACCGGGAGGTACAGACAGATCAGAGAGAGTGAAGTGGGGCA	67720
QY	366	-----	366
Db	67719	TGGTAGGAAGACGTGTGGCCTTGGCTTGGCCCACTGAGAGAGAGCGGGGTAA	67660
QY	366	-----	366
Db	67659	GGAGAAGTAGGCCAGAGTGTGGTCTTGTCCACGTCACACCCCTGCATCTCCTTT	67600
QY	367	-----	367
Db	67599	CTTTCAGCTTAAACAAAGAGGAGGAGGATTCACGCCCATCAGAGGAGGAGAT	67540
QY	385	ALAHLSALAGLYILELGLYARGPHELEUGLYLEUGLSERHISASPISHPROPRO	67481
Db	67539	GF-CACGGTGGGATATTTGGCGGCTTCTTAGGCTTCAGTCCCATGATACCAATCTCC	67421
QY	405	ARGALAAASERARGTHRLEULEUTRPPOLYSARGLUSERLEUHHASGLUGLY	67362
Db	67480	AGGGCAAACTCAAGAGACCAAACTACGTGGCCCAAGAGGAATCCCTTCCACGA-GGC	67422
QY	425	LEUPOLYSASPHISLYSALALALYSGLIAANVALARGLYGLINGLUASPASLYSALA	67444
Db	67421	CTGCCCAAAACCAAGGACCAACCAAGAAAGCTTAGGGGCCAGAGACAAAGAGCC	67362
QY	445	TRPLYSLEULYSALVALASPLAPHELYSSERGLYPROLEUTYRGLINARPROGLYTYR	67302
Db	67361	TGGAAAGCTTAAGGCGTGTGAGCCCTTCAAGTGTGCCCTGTATAGAGGCCAGGCTAC	67302
QY	465	TYRSERIALAPROGLINTHRPROLEUSERPROTHRPROMETPHEPHEPROLEUGLUPROSER	67242
Db	67301	TACAGTGGCCACAGACGCCCTCTAGCCCACTCCCATGTCTTCCCTTAGACATCA	67242

QY	483	AlaProSerLysIleuHisSerValThiThcGlyIleAspThrLysAspLysSerLeuLysThr	504
Db	67241	GGCGCGTAAAGCTTCACAGTGTACAGGATACACCAAAACAAAGCTTTAAAGACT	67182
QY	505	ValSerSerGlyAlaLysLysSerPheGluIleuLeuSerGluSerAspGlyAlaLeuMet	524
Db	67181	GTGAATCTGTGGGGCCAGAAAAGTTTGAATTCCTCTCAGACAGCCATGGGCGCTTGAATG	671222
QY	525	GluHisProGluValSerGlnValArgArgLysThrValGluPheAsnLeuThrAspMet	544
Db	67121	GAGCACCCAGAAATGATCTCAAGTACGAGAGAAAACGTGGAGTTTAACTGACGAGGATATG	67062
QY	545	ProGluIleProGluAsnHisLeuLysGluProLeuGluGlnSer	559
Db	67061	CCAGAGATCCCGAAAATCACTCTCAAGAACCTTTGGAAACAAATCA	67017
RESULT 12			
AK096459			
LOCUS	2500 bp	mRNA	11mer
DEFINITION	Homo sapiens cDNA FLJ39140 fls, clone NTONG2009233, moderately		
ACCESSION	AK096459		
VERSION	AK096459.1	GI:21755962	
KEYWORDS	oligo capping, fls (full insert sequence).		
SOURCE	Homo sapiens tongue cDNA to mRNA, clone_11b:NTONG2 clone:NTONG2009233.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
AUTHORS	1 Oshima,A., Takahashi-Fujii,A., Tanase,T., Inose,N., Takeuchi,K., Arita,M., Mitsuhashi,K., Yuchi,H., Hara,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuno,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Matsutsuma,M., Murakawa,K., Kanohori,K., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahori,K., Masuno,Y., Nagai,K. and Isogai,T.		
TITLE	NEO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2500)		
AUTHORS	Isogai,T. and Yamamoto,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kametari, Kisarazu, Chiba 292-0812, Japan (E-mail:genominfo@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Research Association for Biotechnology (RAB): cDNA full insert sequencing: construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-83'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.		
COMMENT	Location/Qualifiers		
	1..2500		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="NTONG2009233"		
	/issue_type="tongue"		
	/clone_11b="NTONG2"		
	/note="cloning vector: PME18SFL3"		
	229..2235		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="BAC04797.1"		
	/db_xref="GI:21755963"		
	/translation="MTVYSSKVNATFFGGRHLKWKRGSIKLLYRETYVALYT		
	ALISIVRLLLGVGRKPEKISYCDXKADIPPTFVLGGYTVLLVNRMNQGVNLPW		
	PDRMLFLSSSVHSDHGRLRLRTLRKYNVLISLFRSVSTAVYKRPETMDHVEA		
	GMPTDEKRLNRHLKSPHLKWPFIWFGNLAKRNRGRIRSVDSQSLMEENMYR		
	SWCSITFDKDWNGIPIVYTVGVTLVAVTFFACIGRGQFLDPRKGVAGHLDLYIPIF		

```
OY 288 ----- 288
Db 70537 GGGCTGTCTCAGAAAGTGTGTGGTGCCACAAAGTGTGGGGGGCTGAGACCCTAACCTGTC 70478
OY 288 ----- 288
Db 70477 CTTTAAACAGAGTGGTCAAGGAGGAGGCGCTCATGGGGTGTGAAATAGCAGCAGCTGA 70418
OY 288 ----- 288
Db 70417 GGTAAAGGGGAAGCTGGCTTGAAGAGTTCTGCTGAGGGTTTACAGAGCCCTCACCCT 70358
OY 289 ----- 306
      |||ValAlaIglInLeuIleasnProphgIyluAspAspAspPheGlu
Db 70357 GTCCCCAAGGTGGCAGAGCAGCAGCTCAACCCCTTGGAGAGATGATGATTTTGTAG 70298
OY 307 ThrAsnTrpIleValAspArgAsnLeu----- 315
      |||
Db 70297 ACCAACTGAGATTGCGACAGGAATTT--GCAGGTATGGGGAGAGGAGAAACCATACCA 70239
OY 315 ----- 315
Db 70238 TGGACCTTCCCAAGTGAACCAAGAGAGAGACCACACTGTTCTGTAGGAGGCTTCAC 70179
OY 315 ----- 315
Db 70178 AGTGAATGATCAACCTTCCCTCCCTCCCTCTGACAGCAGTCACTCAGAGATTC 70119
OY 315 ----- 315
Db 70118 TCACCTCAATCTTTGAGGCTGACAGGAGCAGCAGCCACTCCCACTTTCACAGGAGGAAA 70059
OY 315 ----- 315
Db 70058 CTGAGGTCCAGAGAGAGAGAGATTCCTCCAAATCATCAGCAGCATACATAGCTCTGC 69999
OY 316 ----- 321
      |||GlnValSerIleuLeuAla
Db 69998 CTGGGATGATCTTCTGCTGGGACTTCTTGTGCTCCGTGACGAGTGTGCCCTGGCT 69939
OY 322 ValAspGluMetHisIleAspLeuProArgMetGluProAspMetTrpAsnLysPro 341
      |||
Db 69938 GTGGATGAGATGTCACAGGAGACCTGCTCGATGGAGGCCGACATGTAAGTAAGGCC 69879
OY 342 GluProGluProProTrpTrpThrAlaIleSerAlaIleGlnPheArgTrpAlaSerPheMetGly 361
      |||
Db 69878 GAGCCACAGCCCCCTACACAGCTGT--TCCGCCAGGTTCCGTGAGCCCTTTATGGGC 69820
OY 362 SerThrPheAsnIle----- 366
      |||
Db 69819 TCCACCTTCAACATCAGGTGTGGCCAGAGCCAGGGGGCTGGGTGGGAAGCCCTCCTAGT 69760
OY 366 ----- 366
Db 69759 GCAGGGGTCTGCCTAGGAACCTAGAAATAGCACTAGTTAATGATACAGGTTGCTTCAGTA 69700
OY 366 ----- 366
Db 69699 AGTGCAGGCACTGTAATGCTCTTTTAAACATTAATAATTTTTCCTCCAAATAT 69640
OY 366 ----- 366
Db 69639 TCTGGTTGTATCCCAAGTTTTCAGATTAATTAAGTAACAGGTTTCAGAGAGTAAGTTG 69580
OY 366 ----- 366
Db 69579 TCCAAGGCCACATAGTACCAAAATGTGATTTGCTACTCGAAGAGACAGCCTGTGATCAG 69520
OY 366 ----- 366
Db 69519 TGATGACGTGAAGCTTAGAGACTGGCTCTTGATCAGAACTATGTTTCTTTCTTT 69460
OY 366 ----- 366

OY 69459 TTGAGACAGTATCTGCTCTGTGCTGCCAGGTTGAGAGCCAGTGGGTGATCTTGCTAC 69400
OY 366 ----- 366
Db 69399 TGCACCTCCGCCCTCTCTGGGTTCAAGTGAATTCCTGCTTACGCCCTCCAGTAGCTGG 69340
OY 366 ----- 366
Db 69339 ATTACAGGTGCCCAACACACACACTGGCTAATTTTGTACTTTAGTAGAGATGAGTTT 69280
OY 366 ----- 366
Db 69279 CACCATGTGGCAGGCTGTCTCCAACTCCTGACAGTAATCTGCCGCTTGCCCTCC 69220
OY 366 ----- 366
Db 69219 CAAATGCTGGAATTAATAGTGTCAAAACATATGTTTCTGATTAAGCTACATGCTTGGAT 69160
OY 366 ----- 366
Db 69159 GGGAAAGTGAAGTGGGGTCTCTGGGATGGGGAGGGGAGCAAAAGTCCAGCAGGAGC 69100
OY 366 ----- 366
Db 69099 CAGGCCATCAGAGTACTCTCTGAATTAATGACTTTGTCTACCGAGTAAGGGCTCAGGCCA 69040
OY 366 ----- 366
Db 69039 CCCACAGACGCCAGACTTAATCCCAATGCTGCCACTTCCCTGATTTCCATCTGAATCCCT 68980
OY 366 ----- 366
Db 68979 CTTGAGCTCAGTGGGCTGAAGGCTATCCACAGCTGTCTCTTCTCCAGAGACACAGA 68920
OY 366 ----- 366
Db 68919 GTTGAAGTGCCTTGAGAGTGTGGGACATGTCAGGGTTCATCAAGGTTTCTTC 68860
OY 366 ----- 366
Db 68859 CACGGTATCAGTGTCTGCTGCTGTGTTCTTTCTTTTAAACGAGATTTC 68800
OY 366 ----- 366
Db 68799 ACTTGTGGCCAGAGCTGAGTGCAGTGGCATATCTCGGCTCAGTCAACTCCGCC 68740
OY 366 ----- 366
Db 68739 TCCAGATTCAAGCAATTCCTGCTCAGCCTCCTGAGTACCTGGGATTTATAGTGCCA 68680
OY 366 ----- 366
Db 68679 GCCACAGCCCGGCTAATTTTGTATTTTATAGAGACAGTTTACACATGTTGGCAG 68620
OY 366 ----- 366
Db 68619 GCTGCTCGAATCTGACCTCAGGTGATCCACCCTCCTCAGCCTCCCAAGTGTGGG 68560
OY 366 ----- 366
Db 68559 ATTACATGTGAGACCACTGTGCCTGCTGCTTTCTTTTAAGAACCAATATCTACT 68500
OY 366 ----- 366
Db 68499 AGACTGAATCAGTTTAAGTACTATAGTATAGATGATGAGAAAGGTTGGAAAGTGA 68440
OY 366 ----- 366
Db 68439 TCAATTAAGGCTGGAGGCTTGTAGGTCAGAAACATTTCTGAGAGTACTTGTAGCC 68380
OY 366 ----- 366
```

Db	72757	GGCCCTGAGGCTCTTCCGAGAGCCTGAGGTGGGTTGCAAGATCTTTTCCAAACAGCAATC	72698
Oy	212	-----	212
Db	72697	CACAGCCCGAGGTGGTCCCTTCTCAGAGAGCCCTCCCTTCTCTCCAACTCTGTGAGTCC	72638
Oy	212	-----	212
Db	72637	TGCTTCCCTTTTGATAGATGAGAAAGTGAGACACAAGAGGTTTATGAGCTTCCATG	72578
Oy	212	-----	212
Db	72577	GCCACACAGCCAGGATGACCAATAGGTACCGAGCCCTGTACTGGAGAAGAGTGGGG	72518
Oy	212	-----	212
Db	72517	GGCAGCCCGGTGGGGGAGGTGTGTTCAGAAACCCCATCCCTCTTCTGCCCCCGAG	72458
Oy	213	GLMERksnthrlleuatgthrlncysgllyhlsleutyralaryasptripleserlle	232
Db	72457	GAGATGAAACCTTGGCTCTACTCAGTGTGAGACACTGATATGCTTACGACTGATTAAGTATC	72398
Oy	233	ProLeuValIyrthr	237
Db	72397	CCACTGGTATATACACAGGTGAGGACTAGGCTGTGAGGCTGCCCTTTTGGGAAACTGAG	72338
Oy	237	-----	237
Db	72337	GCTAGAAAGCAACAAAGAAAGCAAGTGGGGTGGGAAAGGCTACCTTAGAGGCTAAAGTGCCTC	72278
Oy	237	-----	237
Db	72277	CCCTGGAGTTGGGTCCACACTTTGAAGTTGGGTGTGACTTGAAGTGCCTAAAGTCTAA	72218
Oy	237	-----	237
Db	72217	GAGTCCAGGCTCTCGCTGGCCGCCAGTCAGTAAGAGCAATGTATATCCCATATTAA	72158
Oy	237	-----	237
Db	72157	GAGAGTTGGCGGGGTGACAGTGGCTCATGCTGTAAATCCAGACACTTTGGGAAGCTGAGG	72098
Oy	237	-----	237
Db	72097	CAGGTGATCACCTGAGGTGAGAGTTGAGACACAGCCTGGCCACATGCTGAACCCCA	72038
Oy	237	-----	237
Db	72037	TCTCTACTGAAAAATACAGAAATTAAGCTGTGTGTGTGTCACGCTGTAAATCCAGCTACTT	71978
Oy	237	-----	237
Db	71977	GAGAGCTGAGGAGAGAAATCGCTTGAACCCGGGAGGTGAGGTTGCAGTGAGCTGAGA	71918
Oy	237	-----	237
Db	71917	TCATGCCACTGCATCCAGCCTGGGGGACACAGCAAGACTCTGTCTCAAAACAACAACA	71858
Oy	237	-----	237
Db	71857	AACAAACAACAAGGGGTTAACACAGAGCCCTTAAGTCACATTAAGTGTGCAAGTCAAGACA	71798
Oy	237	-----	237
Db	71797	AGGCTTGTCTCTCTCTCAGACTCCAGCCCTGGAGCAATCTGATTTCAAGGTTCCC	71738
Oy	238	-----GlnValValIthrValAlaValIyr	245
Db	71737	ACCTAGCCCTTGTCTACACATCTCTCTCTCTCCCTCCAGGTGTGACTGTGGCGGGTAC	71678
Oy	246	SerPhePheLeuThrCysLeuValIyrlglnPheLeuAsnProAlaIysAlaIyrPro	265
Db	71677	AGCTTCTCCGACTGTCTCTGTGTGGGGGAGGTTTGTGAACCCCAAGGCTTAACCTT	71618

QY	266	GLYHISGLILEUASPLeuValValProValPheThrPheLeuGlnPhePheTyrVal1	285
Db	71617	GGCGATGACCTGGACCTGGTGTGGCCGCTTCACGTTCCGACGTTCTTCGACGTTCTTCTTATGTT	71558
QY	286	gLyTPLeu----- 	288
Db	71557	GGGTGGCTTAAGTGGGGGCTCTCCAGGGGCTGCTGGCGTGGAGCATGGCCAGAGGGGT	71497
QY	288	-----	288
Db	71497	CATGGCCAGCAGCTGCTGAGACGAGATGCAGTGCAGAGAAAGAGGTCTCAGGGGTA	71433
QY	288	-----	288
Db	71437	GAAGACAGCCAGGCGTGTGGCGCACACCTGTAAATCCACACTCTGGAGGCTGAGGCA	71378
QY	288	-----	288
Db	71377	GGAGATGCTTGAACCGGGAGCGAGGTTGTTGATGAGATTCGTCACACTGCAC	71318
QY	288	-----	288
Db	71317	TCCAGCCTGGGCCAAAAGANTGAATCTATCTCAAAAACAACAACAACAACAAG	71258
QY	288	-----	288
Db	71257	CCCTAAGGTTGAGAAGGCCCTGGCCCTTTAGAAGGAGANTGGGACCACTCTCTTATTC	71198
QY	288	-----	288
Db	71197	AAGATGCTGTTGGGCTGTCTTGTTCCTCAGCTCAAGTGGCTTTCAGGTAATCCCTCC	71138
QY	288	-----	288
Db	71137	CACCAAGCCAAATCTCCGAACAGATGTCTGAATCACACAGTTTCTCTCCACCTTTT	71078
QY	288	-----	288
Db	71077	ATCTTCTCTTCTGTGTGTCACCAACCATCTCTCTCTCTTACCTTCTTATTTT	71018
QY	288	-----	288
Db	71017	TGATATGGGGGTAGTCTGTCTGTCTGTCCCTTCTCTGACTGTGACACACACACACA	70958
QY	288	-----	288
Db	70957	CACACACACACACACACATACACACACACACACAGCATTCCTATTCCTAAATTC	70898
QY	288	-----	288
Db	70897	CCCTGCCCCCAGTATCTTTGGTTCTGCAGATCAAAACAATCACATTTTATGCTT	70838
QY	288	-----	288
Db	70837	GAATTCCTCCGGGGTGGCCCAAGTGGCCTGCAGATGTCCCTCGAACCCTAAGCAGACG	70778
QY	288	-----	288
Db	70777	CGTGTACCTTTCGGGGGCTTTGTTAGGCAATTTAGAGTGTCTATCCAGGAATCTGCC	70718
QY	288	-----	288
Db	70717	CACCTAGACTGGCCCTTAGTTGAGCCCACTCAGTATATATCTGTGTCATGAATGAA	70658
QY	288	-----	288
Db	70657	TAAATTAATGCACTCCAGGTAAGATACATGAGTGAAGATAAGCAGTCACTACGCCGA	70598
QY	288	-----	288
Db	70597	GATATCACTCAGGGACAGCTGTGGGTGTTTACGGGAGAGACTGGCTCAGAAGAGTTAAG	70538

ORIGIN

Alignment Scores:

Score: 3.17e-113 Length: 160169
Percent Similarity: 1433.00 Matches: 500
Best Local Similarity: 19.42% Conservative: 0
Query Match: 15.42% Mismatches: 9
DB: 45.93% Indels: 2072
Gaps: 7

US-09-622-964-3 (1-585) x AC051664 (1-160169)

QY 51 ArgLeuAlaLeuThrGlnGlnGlnLeuMetPheGlnLysLeuThrLeuTyrCysAsp 70
Db 74735 AGGCTGGCCCTCAAGGAAACAAACAGCTGATGTTGAGAACTGACTGTATTCGAC 74676
QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
Db 74675 AG-TACATCCAGCTCATCCCAATTCCTTCGTGCTGGTGAATCCCTTCGTGCTGT 74617
QY 82 ----- 82
Db 74616 CCGGTCCTCTGTGGCCGCCAGGCTCCAGACAGGCCAGGAGATCAGAGAGCTGCG 74557
QY 82 ----- 82
Db 74556 GCAGGGGCTGGGAGGGGGGGGAGACGACGAGGAGGTGCGGCTCTCTGTAGGGA 74497
QY 82 ----- 82
Db 74496 AAGGTCGGACTGCACGCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 74437
QY 82 ----- 82
Db 74436 AGCAATGAAMACCCATTTCTGAGGAGAGCGCTGACATCATGTCCTTGAGCCCTCC 74377
QY 82 ----- 82
Db 74376 GCGGGAGGGAGGGGGCTGCGGCAATTCCTGGACCGAGGGGAGCCCGGGTGACAGA 74317
QY 82 ----- 82
Db 74316 ACCCTTGGGGCTCTGCGCGCTCCATGCGAGCTTGCTGCTCTGCTCCGAGCGCT 74257
QY 82 ----- 82
Db 74256 TCCAGAGAGGCTGGGGCTAGCGCCGCTGCAGACAGAAAGCTGAGAGCGAGGATCG 74197
QY 82 ----- 82
Db 74196 CCGGCGTGGGCGCTGGGCTGCGCGAGCCCTGCGCCCTGCGCCCGCCCGCCCGCC 74137
QY 83 -----glyPheTyrValThrLeuValValThrArgTyrPheAsnGlnTyrGlnAsn 99
Db 74136 TCTGCCCGAGCTTCTACGTAGCGCTGCTGCGACCGCCCTGAGAACCACTAGCAGAAC 74077
QY 100 LeuProTyrProAlaArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGln 119
Db 74076 CTGCGCGTGGCCGAGCCGCTGATGAGCTGTGCTGGGCTTCGTCGAGAGGCAAGAGAG 74017
QY 120 GlnSerArgLeuLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlnLysValLeuIle 139
Db 74016 CAAAGCGGGGCTGCGGGCGAGCGCTATCCGCTACAGCAACCTGGGCAAGCTCTCATC 73957
QY 140 LeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGln 159
Db 73956 CTGCGAGCTGACACACCGAGTCTACAGAGGCTTCCCGAGCGCCAGCACTGTGTCAA 73897
QY 160 Ala----- 160
Db 73896 GC-AGTGGGCGAGCCGAGCAACGGGGAGGACCCGAGGAGAGCCAGGGGCCGAGATGG 73838
QY 160 ----- 160

Db 73837 GCGCGAGCAAGTAATGAATGGGTGGAGCAAAAGTCCCGGAGCTCGGGGATGGGTG 73778
QY 160 ----- 160
Db 73777 GAGCCAGAGTGGGTGTAGTCAAGATTGGGGTCCAAATGGCGGAGAGACTCGGGT 73718
QY 160 ----- 160
Db 73717 GTCTGAAGTGGGGCGAGGCCAGAGGCCACCTCCGAGATGAGACTGTAGGCGAGGC 73658
QY 160 ----- 160
Db 73657 TAAAGACCTTGAGGGATATGAAAGAGGTGAGCGGCTTGGGAAGTGTGATCTA 73598
QY 160 ----- 160
Db 73597 GGGTCTACTCCCTCTGCGCTTCCCTCTTGATCTCCGGTTCCACTGTGAGGATAGG 73538
QY 160 ----- 160
Db 73537 GACATTGTCTCTGACACCCCTCAGCCTGGGCTGACTGTCTGTATTAAGACAGA 73478
QY 160 ----- 160
Db 73477 CCCAGCTAGCGGTGGTGGCTGTGCGCTGTAAATCCAGTGTCTTAGAGGCAAGGTGG 73418
QY 160 ----- 160
Db 73417 AAGATCGCTTGAAGCCAGCTGTTTGAAGCCCGCTGAGCAATATGACGAGCCCATCT 73358
QY 160 ----- 160
Db 73357 CTACAAAACATTAAATTTAGCAGGGGCAATGTGGCGTGTGCTGCTGAGTGAAGT 73298
QY 160 ----- 160
Db 73297 TATCGGAGGCTGAGAGGAGAGAGATCTTGAAGCCAGATTCAGAGCTGAGTGGCT 73238
QY 160 ----- 160
Db 73237 AAGATGCAAGCTGCTCACTCCAACTCGGTGACAGGCAAGCCCTTCTGTGAATAA 73178
QY 160 ----- 160
Db 73177 ATAAATACCTGCGCCACATGTCAAGCCAGAAACAGACCTAGTAGTGTCTAGAAATTT 73118
QY 160 ----- 160
Db 73117 TTTGTTGTGAAGAAAGAGATGGCAAGAGTGTAGTTCTCTATAGGTCAAGAGT 73058
QY 161 -----gly 161
Db 73057 GCGGCGATCCCTTCTGAGGTTCTCCACCCAGCGCTTCTTCACTGCTCACTGCAAGC 72998
QY 162 PheMetThrProAlaGlnHisLysGlnLeuLysLysLeuSerLeuProHisAsnMetPhe 181
Db 72997 TTTTGAAGCTCGGAGAAACACAAACAGTGTGAGAACTGAGCCCTAACACACATGTTTC 72938
QY 182 TrpValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrPheGlnGlyArgIle 201
Db 72937 TGGGTGCGCTGGGGTGTGTTTCCAACTGTCAATGAAGCGTGGCTTGGAGGTGAAATTC 72878
QY 202 ArgAspProIleLeuLeuGlnSerLeuLeuAsn----- 212
Db 72877 CCGGACCTATCTGCTGTCAGAGCTGTCAAGCTGAGCCACTGTACAGACAGGCTGC 72818
QY 212 ----- 212
Db 72817 CGCAGATGGGAAGGGCTGTGTCTCACAGGAAACAGGTTTCTTACAAGAGAGCCTTG 72758
QY 212 ----- 212

REFERENCE 1 (bases 1 to 160169)
 AUTHORS Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 160169)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (15-APR-2000) Genome Sequencing Center, Washington
 MO 63108 USA
 On Apr 23, 2000 this sequence version replaced g1.7574984.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H_NH0810P12
 ----- Summary Statistics -----
 Sequencing vector: M13: 100%
 Sequencing vector: plasmid: 0%
 Chemistry: Dye-terminator Big Dye: 0% of reads
 Assembly program: Phrap: version 0.990319
 Consensus quality: 147316 bases at least Q40
 Consensus quality: 151042 bases at least Q30
 Consensus quality: 152875 bases at least Q20
 Insert size: 16500: agarose-rp
 Insert size: 157869: sum-of-contigs
 Quality coverage: 3.91 in Q20 bases; agarose-rp
 Quality coverage: 4.17 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 24 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1926: contig of 1926 bp in length
* 1927 2025: gap of unknown length
* 2027 4202: contig of 2176 bp in length
* 4203 4302: gap of unknown length
* 4303 7163: contig of 2861 bp in length
* 7164 7263: gap of unknown length
* 7264 9961: contig of 2698 bp in length
* 9962 10061: gap of unknown length
* 10062 12575: contig of 2514 bp in length
* 12576 12675: gap of unknown length
* 12676 15545: contig of 2870 bp in length
* 15546 15645: gap of unknown length
* 15646 20941: contig of 5296 bp in length
* 20942 21041: gap of unknown length
* 21042 23966: contig of 2925 bp in length
* 23967 24066: gap of unknown length
* 24067 27231: contig of 3165 bp in length
* 27232 27332: gap of unknown length
* 27333 30550: contig of 3219 bp in length
* 30551 30650: gap of unknown length
* 30651 35202: contig of 4552 bp in length
* 35203 35302: gap of unknown length
* 35303 40201: contig of 4899 bp in length
* 40202 40301: gap of unknown length
* 40302 45758: contig of 5457 bp in length
* 45759 45858: gap of unknown length
* 45859 51249: contig of 5391 bp in length
* 51250 51349: gap of unknown length
* 51350 58267: contig of 6918 bp in length
* 58268 66916: gap of unknown length
* 66917 67016: contig of 8549 bp in length
* 67017 77145: gap of unknown length
* 77145: contig of 10129 bp in length

```

FEATURES

source	1. 160169	Location/Qualifiers
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="11"	
	/clone="RP11-810P12"	
	1. 1926	
misc_feature	/note="assembly_name:Contig18"	2027..4202
misc_feature	/note="assembly_name:Contig19"	4303..7163
misc_feature	/note="assembly_name:Contig20"	clone_end:77
	vector_side:left"	7264..9961
misc_feature	/note="assembly_name:Contig21"	10062..12575
misc_feature	/note="assembly_name:Contig22"	12676..15545
misc_feature	/note="assembly_name:Contig23"	15646..20941
misc_feature	/note="assembly_name:Contig24"	21042..23966
misc_feature	/note="assembly_name:Contig25"	24067..27231
misc_feature	/note="assembly_name:Contig26"	27332..30550
misc_feature	/note="assembly_name:Contig27"	30651..35202
misc_feature	/note="assembly_name:Contig28"	35303..40201
misc_feature	/note="assembly_name:Contig29"	40302..45758
misc_feature	/note="assembly_name:Contig30"	45859..51249
misc_feature	/note="assembly_name:Contig31"	51350..58267
misc_feature	/note="assembly_name:Contig32"	58368..66916
misc_feature	/note="assembly_name:Contig33"	67017..77145
misc_feature	/note="assembly_name:Contig34"	77246..86493
misc_feature	/note="assembly_name:Contig35"	86594..96688
misc_feature	/note="assembly_name:Contig36"	96789..107169
misc_feature	/note="assembly_name:Contig37"	107270..117047
misc_feature	/note="assembly_name:Contig38"	117148..127120
misc_feature	/note="assembly_name:Contig39"	clone_end:sp6
	vector_side:right"	127221..142029
misc_feature	/note="assembly_name:Contig40"	142130..160169
misc_feature	/note="assembly_name:Contig41"	36693 c 37296 g 42226 t 2307 others

BASE COUNT 41647 a 36693 c 37296 g 42226 t 2307 others

Dh 90977 GGACACAGATTGTAAGTGCCTTGAGAGTGTGGGACATGTCAAGGTTCAATCA 91036
Qy 366 ----- 366
Dh 91037 GGGTTTCTCCAGCGATTCAGTGTCTGCTGCTGCTCTTTCTTTTCTTTTAA 91096
Qy 366 ----- 366
Dh 91097 ACGGAGTTTCACTTGTGTGCCAGAGCTGAGTGCAGTGGCATATCTGGCTCACTGC 91156
Qy 366 ----- 366
Dh 91157 AACCTCCGCTCCAGATTCAGCAATTCCTGCTCAGCTCCCTGATAGTGGGATT 91216
Qy 366 ----- 366
Dh 91217 ATAGTGCACGACCAAGCCCGGCTAATTTTGTATTTTATAGACAGTTTCACCA 91276
Qy 366 ----- 366
Dh 91277 TGTGGCCAGGCTGTCTGAGACTCCTGACCTGAGTATCCACCTCCAGCTCCCA 91336
Qy 366 ----- 366
Dh 91337 AAGTCTGGGATTACATGTGTGAGCCAGCTGCTGCTGCTGCTTTTAAAGAACCA 91396
Qy 366 ----- 366
Dh 91397 ATATCTACTAGACTGCAATCGAGTTTACTACAGTCTATAGATGTGAGAGAGGTT 91456
Qy 366 ----- 366
Dh 91457 GGGAGGATGATCAATGAAGCTGAGGCTGCTTGTAGTCAAGAACTTCTGAGAGATG 91516
Qy 366 ----- 366
Dh 91517 ACTTTAGAGCCCTACATNGTCTGTACCCAGCAGCTGAAGGTTGTGAGGGGTGGAGGG 91576
Qy 366 ----- 366
Dh 91577 CTGAACAGACAGATTAAGCATAGACCTTGTCTCAAGAGATGCACATTTATGAGGG 91636
Qy 366 ----- 366
Dh 91637 AGCTCAAAACCAAGTCTCAAACTGTGATACAGGTACAAAGTACTGATGTCCAGAAA 91696
Qy 366 ----- 366
Dh 91697 GGGACAGAACATGGAACACAGTCATCTTGTCTGCTGGAGGGCGCTTCCAGCTGGGTC 91756
Qy 366 ----- 366
Dh 91757 TGAAGCTGAGCCATGGAACATGGAAGATCTGAACCTTGGGCAAGGGACGCCATCTCT 91816
Qy 366 ----- 366
Dh 91817 CTGATAGATAGCTTTCCCTTGCAGGGTAAAGTCTGGGGTCCCGGATGCTGTGCT 91876
Qy 366 ----- 366
Dh 91877 AGGAAGTCAATTTCTTGTGTGATGTCACTCCAGTGTGAACACCAAAATTCCTGGCAT 91936
Qy 366 ----- 366
Dh 91937 TGCCCAAGTCACTCATGAGGCCCTATCTGAACACATCATGCCAGGACACAGTGTCTG 91996
Qy 366 ----- 366
Dh 91997 ACTGCTGAGAGTGAGGGGTTTACAGGGAAGTGAATGAGAGGCTTTACAGGCCCA 92056
Qy 366 ----- 366
Dh 92057 GCGGGGGTGGTGGGGGTTGGATGTTAACTGTGTCAGAGGGATCAACAACAGTGC 92116

Qy 366 ----- 366
Dh 92117 AGTGAGCTGGGCTGTGAGGATCACCAGGATACAGACAGATCAGAGAGAGTGAG 92176
Qy 366 ----- 366
Dh 92177 AGCTGGGCGATGGTGAGAGAGACGATGTGGCTTGGCTTGGGCAACTGAGAGAGAG 92236
Qy 366 ----- 366
Dh 92237 CGGGGCTAAGGGAAGTAAGCCAGGTGTGTCTTGTCCACTGGCTCAGCCCTGCA 92296
Qy 367 ----- SerLeuasnlysgluglumetgluphegluprobangluas 381
Dh 92297 TCTCTGTGTTCTTCCAGGCTGAACAAAGAGAGATGAGATTCACACCCATATAGAGGA 92356
Qy 381 pglugluaspralalhisalgllylellellyarphleuglyleuglnserhisaphl 401
Dh 92357 CGAGAGAGATGCTCAGCGTGGCATTCATGGCCCTTCTTCAAGGCTCAGTCCATGATCA 92416
Qy 401 shisproaratgalaasnserarqthrlyleuleutpprolyarigluiserleu 421
Dh 92417 CCATCTCCAGGGCAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCT 92476
Qy 421 uhlsgluglyleuprolyasnhislysalalalysglnasnlarglyglngluas 441
Dh 92477 CCACGAGGCTGCCCCAAAACACACAGCCACAAACAGAAAGTTAGGGGCCAGAGAGA 92536
Qy 441 pasnlysalatrplyleuleylalavalaspralpheylusserglyproleutyglinar 461
Dh 92537 CAACAAGGCTGTGAAGCTTAAGGCTGTGGAGGCTTCAAGTGTGCCCTACTATCAGAG 92596
Qy 461 pProcllytyrlyseralaproglthrproleuserprothrprometphepserpro 481
Dh 92597 GCCAGGCTACTACAGTGCCTCCACAGACGCCCCCTCAGCCCATCCCATTTTCCCT 92656
Qy 481 ugluproseralaproserlyleuleylhiservalthrlyleasprlyleasprlyse 501
Dh 92657 AGAACCATCAGGCGCGTCAAAACCTTCAACAGTGTCCAGGCAATGACCAACAAAGCAAAAG 92716
Qy 501 rleulsthrlyalserserglyalalyserhegluleuleusergluserpogl 521
Dh 92717 CTTAAAGACTGTGATTTCTGGGCCCAAGAAAAGTTTGAATTCCTCAGAGAGGATGG 92776
Qy 521 yAlaleumetcluhisprogluvalasergluvalarghrlysthrlyalglupheasnl 541
Dh 92777 GGCCTGTATGAGCACCCAGAAATGTCAAGTGAAGGAGAAACCGTGAATTAACCT 92836
Qy 541 uThrasmhetprogluileprogluasnhisleuleylusgluproleugluinserproth 561
Dh 92837 GACGGATATGCCAGAGATCCCGGAAATCACCTCAAAACCTTGGTGAACATCACCAAC 92896
Qy 561 rasnllhissthrlyleuleylasprhismetasprprotyrtrpialeugluasnaras 581
Dh 92897 CAACATACACACTACATCAAGATACATGATTCATTATGGGCTTGGCAAAACAGGTC 92956
Qy 581 pglualalhis 584
Dh 92957 TGTCTCTCAC 92966

RESULT 11
AC051664/c 160169 bp DNA linear HTG 01-SEP-2000
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-810P12, WORKING DRAFT
SEQUENCE, 24 unordered pieces.
AC051664
AC051664.2 GI:7637836
VERSION
KEYWORDS HTG: HTGS_PHASE1, HTGS_DRAFT.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

QY 288 ----- 288
Db 88817 AGTTTCTCTCACCCTTTAATCTTCTCTCTCTGTGCCACCACCTCTCTCTCTCT 88876
QY 288 ----- 288
Db 88877 TCTCTACCTTCTTATTTTGGTAATGGGGGTGTAACTCTCTCTCTCTCTCTCTCT 88936
QY 288 ----- 288
Db 88937 CACTGTGACACACACACACACACACACACACACACACACACATTCCTATTC 88996
QY 288 ----- 288
Db 88997 TCTAATTTCCCCCTGCCCCCAGTATCTTTGGTTCTGCGATCAAAACAAATCACAC 89056
QY 288 ----- 288
Db 89057 TTTTATGCTTGAATTCCTCAGGGTGCCCCAGTGCCCGCAAGATGTCCCTGGACCCCT 89116
QY 288 ----- 288
Db 89117 AAGCAGACGGGTGTACCTCTTCGGGGCTTTGTAGGGCATTTTAGAGTGTCTATCCA 89176
QY 288 ----- 288
Db 89177 GGAATCTGCCACCTAGACTGCCCTTATGTTACGCCAGCTTCAGTATATATCTGTGTG 89236
QY 288 ----- 288
Db 89237 CATGAATGAATAAATATATGCAACTCCAGGTAAAGATATAGAGTGAATAAAGCAGTG 89296
QY 288 ----- 288
Db 89297 ACTCAGCCGAGTATACACTCAGGGAGACGCTGTGGGTGTTCAGGGAAGAGTGGCTCAGA 89356
QY 288 ----- 288
Db 89357 AGAGTTAGAGGGGCTGTGTCCAGAAAGTGTGGTGGCCACAACTGTGGGGGCTGGAGC 89416
QY 288 ----- 288
Db 89417 CCTAAACTCTGCTTTGAGACAGAGTGTCTCAGCAGCAAGAGGCTTCATGGGGTGTGAAT 89476
QY 288 ----- 288
Db 89477 AGCAGCAGCTGAGGTTTAAAGGGGGAAGCTGGCTTTGAGAGTCTGCTGAGGGTTTAC 89536
QY 289 ----- 289
Db 89537 AGAGCTCACCCTGTCCCAAGGGGAGAGCAGTCTATCAACCCCTTTTGGAGAGATGAT 89596
QY 303 AspAspPheGluThrAsnTrpIleValAspArgAsnLeu----- 315
Db 89597 GATGATTTTGGACCAACTGTGATGTCTGACAGGAATTTGACAGTATGGGAGAGAGAGAG 89656
QY 315 ----- 315
Db 89657 AAACCATACCATGAGACTTCCCAAAAGTGGACCCAAAGAGAGAGACCCACTGTTCTAG 89716
QY 315 ----- 315
Db 89717 GGAGCCCTCAGAGTGAATGATCAACCTTCCCTCTCTCTCTGAGCAGCATTCATTCAC 89776
QY 315 ----- 315
Db 89777 TCACAGGATTCACCTCAATCTTTGAGGCTGACAGCAGGACACCATCTCCCATTTTCAC 89836
QY 315 ----- 315
Db 89837 AGGCAAGGAAGTGAAGGTCCAGAGAGAGGAGATTCCTCCAACTCATCAGCAGCATATAC 89896
QY 316 ----- 316
Db 316 ----- 316

Db 89897 AAGTCTCTCGGGAGATGATCTTCTGTGGAGACTTCTTGTGCTCCCTGGTGAGCAGAGTGC 89956
QY 318 rIeuLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTrp 338
Db 89957 CCGTGGCTGTGATGAGATGACACAGAGACCTGCTCGATGAGGCGGAGCATGTACTAG 90016
QY 338 PAsnLysProGluProGlnProProTrpThrAlaIleSerAlaGlnPheArgArgAlaSe 358
Db 90017 GAATAGCCCGAGCCAGCCAGCCCGCTTACACAGCTCTTCCGCCAGTTCCGTGAGCCCTC 90076
QY 358 rPheMetLysSerThrPheAsnIle----- 366
Db 90077 CTTTATGGGCTCCACCTTCAACATCAGGTGTGGCCAGAGCCAGGGGCTGGGTGGAGAC 90136
QY 366 ----- 366
Db 90137 CCTCTCTAGTCAGAGGGTCTGCTTAGAAGACTTAGAATAGACTAGTATAGCATACAGT 90196
QY 366 ----- 366
Db 90197 TGTCTAGTAGTGTAGGACACTGATATGCTCTTTAATACATTAATATTTTTC 90256
QY 366 ----- 366
Db 90257 TCCCAATATTCGTGTTGTATTCACCAATTTTCAGATATTAAGTACAGGTTTCAGAGA 90316
QY 366 ----- 366
Db 90317 GAGTAAGTGTCCAGGCCACATAGCTACCAAAATGTGATTTGCTACTCGAAGGACAGC 90376
QY 366 ----- 366
Db 90377 CTATGATCAGTATGATGAGTGAAGTGAAGCTGTGCTCTGTCTATCCAGAACTATGTT 90436
QY 366 ----- 366
Db 90437 TCTTTCTTTTGTAGACAGTATCTGCTGTGTGCCAGGTGGAGGCAAGTGGCTGAT 90496
QY 366 ----- 366
Db 90497 CTTGGCTACTGCAACTCCGCCCTCTGCGTTCAAGTATCTCTGCTTACGCTCCCC 90556
QY 366 ----- 366
Db 90557 AGTAGCTGGATTACAGGTGCCACACACAACTGGCTAATTTTGTACTTTAGTAGA 90616
QY 366 ----- 366
Db 90617 GATGAGTTTCCACCATGTGTCAGAGCTGTCTCCAACTCCTGACAGTATCTGCCGC 90676
QY 366 ----- 366
Db 90677 TTTGGCTCCCAAAATGCTGAATATAGTGTCAAACTATGTCTGTGTAAGTACG 90736
QY 366 ----- 366
Db 90737 ATGCTTGATGGAGTGAAGTGGGTTCCCTGGATGGGGAGGGGAGCAAAAGTCCC 90796
QY 366 ----- 366
Db 90797 AGCAGCAGCCAGGCAATCACAGGTACCTCTGTAATGACTTGTCTTACCGAGTAAGG 90856
QY 366 ----- 366
Db 90857 GCTCAGGCCAACCACAGCAGGCACTATACCCACATGATGCCACTTCCCTGATTCAT 90916
QY 366 ----- 366
Db 90917 CTGAACTCTTGTAGCTCAGTGGCTGAAGGGCTATCCACAGTGTCTTCTCCCA 90976
QY 366 ----- 366
Db 366 ----- 366

Db 86598 GTATGGAGGCTGAGGAGGAGATCACTTGAGCCACAGATTCAGGCTGCACTGGCC 86657
QY 159 ----- 159
Db 86658 TAAAGTGCACCGCTGCATCTCCAACTCGGTGACAGAGCCAGACCTTTCTCTGGAATA 86717
QY 159 ----- 159
Db 86718 AATAATACCTGCCACATGCTCAGCCAGACAGCCTAGTAGTGTCTCAGAAATTT 86777
QY 159 ----- 159
Db 86778 TTTGTGTGTAAGAAAGAGAGATGCAGAGAGTGTGAGTTCATATAGTCAGCAG 86837
QY 160 ----- 160
Db 86838 TGCCGGCCATCCCTTCTGAGGTCTCCACACCTCCTCTTCACTCCACTGCGAGG 86897
QY 161 yphemeThrProAlaIuHlslsLysGlnLeuGluLysLeuSerLeuProHIsasmetPh 161
Db 86898 CTTATGACTCCGCGACAGACAGAGTGGAGAACTGAGCCTACACACAACTGT 86957
QY 181 eTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyGlyArg 201
Db 86958 CTGGGTGCTCTGGGTGTGGTGTGCTGCAACCTGTCAATGAAAGCGTGGCTGAGGTGAA 87017
QY 201 eArGAspProIleLeuLeuGlnSerLeuLeuAsn ----- 212
Db 87018 CCGGACCCATCTCTGCTCAGAGCCTGCGTGAACGTGAGCCACTGACAGACAGGCTG 87077
QY 212 ----- 212
Db 87078 CCGCAGATGGGAAGGCTTGTGTGTCACAGAAACAAGTTCTTCAAAAGAGACCTT 87137
QY 212 ----- 212
Db 87138 GGGCCCTGAGGCTTCCGAGAGCCTGAGGTGGGTTGACAGATCTTTCCAAACAGCA 87197
QY 212 ----- 212
Db 87198 TCCACAGCCGAGGTGCTCCTTCAGAGCCCTCCTCTTCCAGTCTGTAGGT 87257
QY 212 ----- 212
Db 87258 CCGTGTCCCTTTTATAGATGAGAGAGCTGAGACACAAAGATTAGTGAAGTCCCA 87317
QY 212 ----- 212
Db 87318 TGGCACACAGCCAGGAATGACATAGTACAGAGCCCTGTACTGGAGAAAGGTGG 87377
QY 212 ----- 212
Db 87378 GGGCGAGCCAGGAGTGGGGAGGTGTTCAGAACCCATCCCTCTCTGCCCCC 87437
QY 213 --GlumeLAsnThrLeuArgThrGlnCysGlyHIsLeuTrpAlaTrpAspTrpIleSer 232
Db 87438 AGGAGATTAACACCTTGCTGCTCAGTGTGAGACCTGTATGCTTACAGCTGATTAATA 87497
QY 232 leProLeuValLysTrpGlnVal ----- 239
Db 87498 TCCCACTGTGTATACACAGGT -GAGGACTAGGCTGTGAGGCTGCCCTTTGGGAACT 87556
QY 239 ----- 239
Db 87557 GAGGCTGAAGAGCAAGAGAGAGCTGGGTGGAGGCTCAGTAGAGGCTAAGTGG 87616
QY 239 ----- 239
Db 87617 CTCCTCGGAGATTGGTCCACACTTTGAAGTTGGGTCTGACTTTGAAGTCCCAAGTTC 87676
QY 239 ----- 239
Db 87677 TAAAGTCCAGGCTCTGCTGGCCAGTCCAGTAGAGGCAATGTGATTTATCCCAATTT 87736

QY 239 ----- 239
Db 87737 AAAGAGAGTTGGCCGGGACAGAGTGGCTCATGCTGTATATCCAGACCTTTGGGAAGCTG 87796
QY 239 ----- 239
Db 87797 AGCGAGTGTATCAGCTGAGGTGAGAGTTGAGAGCCAGCCTGGCAACATGTGTAAAC 87856
QY 239 ----- 239
Db 87857 CCATCTTACTGAAATATACGAATTAAGTGTGTGTGTGTGTCACGCTGTATATCCAGCTA 87916
QY 239 ----- 239
Db 87917 CTGGGAGGCTGAGGCGAGGAATCGTTGAAACCCGGAGGTGAGGTTGACGTGAGCTG 87976
QY 239 ----- 239
Db 87977 AGATCATGCCACTGCACCTCCAGCCTGGGCGACACAGACAGACTGTCTCAACAAACAA 88036
QY 239 ----- 239
Db 88037 ACAACAAACAAACAAACAAACAAAGGGTTAAACAGCCCTAAGTCAAGTAAAGTGC 88096
QY 239 ----- 239
Db 88097 AAGTCAGAACAAAGGCTGTGTCTGTCTCAGACTCCAGCCCTGAGACTCTGATTT 88156
QY 239 ----- 239
Db 88157 TCAAGGTTCCCACTAGCCCTTTGCTACACACTCTCTCTCTCTCTCTCTCTCTCCAG 88216
QY 240 --ValThrValAlaValLysSerPhePheLeuThrCysLeuValGlyArgGlnPheLeu 259
Db 88217 TGTGACTGTGGCGGTGATACACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88276
QY 259 snProAlaLysAlaTrpProGlyHIsGlnLeuAspLeuValProValPheTrpPhe 279
Db 88277 ACCCAGCCAAAGGCTTACCTGTGCGCATGTAGAGCTGTGCTGTGCGGCTTACAGTTTC 88336
QY 279 euGlnPhePhePheTrpValGlyTrpLeu ----- 288
Db 88337 TGCAGTTCTTCTTATGT 88396
QY 288 ----- 288
Db 88397 GGAGGCTGGCCAGAGGGGTCTATGGCCAGCACTGCTGAGACAGATGCCAGTGCAGG 88456
QY 288 ----- 288
Db 88457 AAAGAGGTCTCAGGGGTAGAAAGACAGCAGGCGTGTGTGCACACACACTGTAAATCCAG 88516
QY 288 ----- 288
Db 88517 CTACTCGGAGGCTGAGGCGAGAGAAATCGCTTGAACCGGAGCGAGTTGTGTGAG 88576
QY 288 ----- 288
Db 88577 TTGAGATCGTGCACATGCTCAGCTCAGCCTGGGCAAAAGATGAATCTATCTCAAAAA 88636
QY 288 ----- 288
Db 88637 ACAACAAACAAACAAACAAACCTTAAGTTTCAGAAAGCCCTGCTTTAGAGAGCAGAGC 88696
QY 288 ----- 288
Db 88697 GGCACACCTCTCTTATTCAGATGCTGTGTGGCTGTGTCTTGTCTCTCACTCAAGTGC 88756
QY 288 ----- 288
Db 88757 CTGTGCAAGTATTCCTCCACACAGCAATCTCCGAACAGATGTTCTGAATCAGAC 88816

```

FEATURES          * 113520 133683: contig of 20164 bp in length.
SOURCE
1. 133683
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11"
   /clone="CTD-3231N5"
   /clone_11b="CTD Human BAC"
misc_feature
1. 1459
   /note="assembly_fragment"
misc_feature
1560. 2520
   /note="assembly_fragment"
misc_feature
2621. 6334
   /note="assembly_fragment"
misc_feature
6435. 11913
   /note="assembly_fragment"
misc_feature
12014. 20756
   /note="assembly_fragment"
misc_feature
20857. 29735
   /note="assembly_fragment"
misc_feature
29836. 38645
   /note="assembly_fragment"
misc_feature
38746. 52193
   /note="assembly_fragment"
misc_feature
52294. 65209
   /note="assembly_fragment"
misc_feature
65310. 79796
   /note="assembly_fragment"
misc_feature
79897. 95896
   /note="assembly_fragment"
misc_feature
95997. 113419
   /note="assembly_fragment"
misc_feature
113520. 133683
   /note="assembly_fragment"
misc_feature
113520 a 34761 c 35336 g 30857 t 1209 others
ORIGIN
BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others
ALIGNMENT SCORES:
Pred. No.: 3.89e-131 Length: 133683
Score: 1636.50 Matches: 528
Percent Similarity: 20.26% Conservative: 0
Best Local Similarity: 20.26% Mismatches: 6
Query Match: 52.45% Indels: 2073
DB: 2 Gaps: 7
US-09-622-964-3 (1-585) x AC084857 (1-133683)
QY 51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluLysLeuThrLeuTyrCysasp 70
Db 85158 AGGCTGGCCCTCAGCGAAGAACTGATGTTGAGAACTGACTCTGATTGGCAGC 85217
QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu 82
Db 85218 AGCTACATACAGCTATCCCATTTCTGCTGGGTGAGTACCCCTCTGCTGTT 85217
QY 82 82 82
Db 85278 CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGGCCAGGGAGGATGACGAGAGCTGCG 85337
QY 82 82 82
Db 85338 GCAAGGGGCTGGGGAGGGGGGGGGAACCCAGCGGAGGTCGGCGCTCTGTAGGG 85397
QY 82 82 82
Db 85398 AAAGTGGGAGCTGACAGCAGAGAACTGAAGTTAGAGTTAGAGTCTGCGCGT 85457
QY 82 82 82
Db 85458 TAGCAATGAAACCCATTTTCTGAGGGAGCGCTGACATCATGCTCCGAGACCCCTG 85517
QY 82 82 82

```

```

Db 85518 CGCGGAGGGAGGGGGGTCTGGCGGATTTCTGGACACAGGAGGAGCCCCGGGTGACA 85577
QY 82 82 82
Db 85578 GAACCTTGGGGCTCTCGCGCTCCATCGAGAGGCTCTGCTCTCTCCCTCCAGCGC 85637
QY 82 82 82
Db 85638 CTTCAGAGAGGCTGGGGGCTAGGCCGCTCGCAGACAGAAAGCTGGAGAGCGAGCAT 85697
QY 82 82 82
Db 85698 CGCGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCT 85757
QY 83 83 83
Db 85758 CCTCTCTGCCAGGCTTCTACGTGACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 85817
QY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluLysAsp 118
Db 85818 AACCTGGCGTGGCCGACCGCTCATGAGCTGGTGGGCTGCTGCGCAAGCAAGAC 85877
QY 119 GluGlnSerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
Db 85878 GAGCAAGGCGGCTGCTGCGCGCGCGCTCATCCCTACGCGCACCTGGGCAACGCTGC 85937
QY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnIleuVal 158
Db 85938 ATCTGGCGCAGCGTACGACCGCAGCTACAAAGCCTTCCCGAGCGCCAGCAAGCTGGTG 85997
QY 159 Gln 159
Db 85998 CAAGCAGGTGGCGGAGCGGAGCAGCAAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGT 86057
QY 159 159 159
Db 86058 GGGCGGCGAGAAATGAAGATGGGTGAGCCAAAGTCCCGGAGCTGGGGATTGGGT 86117
QY 159 159 159
Db 86118 GGAGCGAGAGATGGGGTGTGTGCAAGATTTGGGGTCCAAATTTGGCGGAGACGACTCGG 86177
QY 159 159 159
Db 86178 TGTCTGAAGTGGGGCGAGGCCAGGACCCACCTCCGAGAGTAGAGTCTGAGGCGAGG 86237
QY 159 159 159
Db 86238 CTAAGGACCTTGAAGGATTAATGAAAGAGGTTGACGGCTTGGGATCTGTGAGGTACT 86297
QY 159 159 159
Db 86298 AGGCTACTTCCCTGCGCTTGCCTTGTGATCTCGGTTCCACTCTGAGAGTATG 86357
QY 159 159 159
Db 86358 GGACATTGGTCTGACACCCCTGAGCTGCTGACCTGCTGCTGTTAATAGACAG 86417
QY 159 159 159
Db 86418 ACCGAGCTAGCGCTGTGGCTTCGCTGTAAATCCAGTCTTTAGAGGCAAAAGTGG 86477
QY 159 159 159
Db 86478 GAAGATGCTTAGACCCAGGTGTTGAGAGCCCTGAGCAACATFAGGAGACCCCATC 86537
QY 159 159 159
Db 86538 TCTACAAAACATTAAATTAGCAGGGCATGTGGGCTGTGCTGTAGTCTGAGGCTGA 86597
QY 159 159 159

```

Db	12418	AAGGAGAGTAAGGCGAGTGTGCTTGTCCACTGGCTGACCCCTGCATCTCTG	12477
Qy	367	-----SerLeuAsnLysGluGluMetGluPheGlnProAsnGlnGluAspGlu	383
Db	12478	TTTCTTCCACCCCGAAGAGAGAGATGATTCAGCCCAATCAGAGAGAGAGAG	12577
Qy	384	AspAlaHisAlaGlyLeuLeuGlyArgPheLeuGlnGlnSerHisAspHisPro	403
Db	12538	GATGCTCAGCTGATGATCATTTGGCCGCTTCTAGGCTGCGAGTCCCATGATCACC	12597
Qy	404	ProArgAlaAsnSerArgThrLysLeuLeuTrpProLysArgGluSerLeuHisGlu	423
Db	12598	CCGAGGCGCAACTCAGAGACCAACTGCTGGCCCAAGAGAGAGATCCCTCTCAGAG	12657
Qy	424	GlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGluAspAsnLys	443
Db	12658	GGCCTGCCCAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	12717
Qy	444	AlaTrpLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTrpGlnArgProGly	463
Db	12718	GCCAGGAAGCTTAAGCTGTGAGAGCCTTCAAGTCTGCCCACTGTATCAGAGAGAGC	12777
Qy	464	TyrTrpSerAlaProGlnThrProLeuSerProThrPheMetPhePheProLeuGluPro	483
Db	12778	TACTACAGTGCACAG	12837
Qy	484	SerAlaProSerLysLeuHisSerValThrGlyLysAspThrLysAspLysSerLeuLys	503
Db	12838	TCAGCCCGCTCAAG	12897
Qy	504	ThrValSerSerGlyLysLysLysSerPheGluLeuLeuSerGluSerAspGlyAlaLeu	523
Db	12898	ACTGTGAGTTCGCGGCGCAAGAAAGTTTGAATTGCTCTCAGAGAGAGAGAGAGAG	12957
Qy	524	MetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsnLeuThrAsp	543
Db	12958	ATGGAGACCCAGAAATATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	13017
Qy	544	MetProGluLeuProGluAsnHisLysLysGluProLeuGluGlnSerProThrAsnLys	563
Db	13018	ATGGCAGAGATCCCGCAAAATACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	13077
Qy	564	HisThrThrLeuLysAspHisMetAspProTyrTyrPalaLeuGlnAsnArgAspGluAla	583
Db	13078	CACACTACACTCAAGATCATCATGATCTTATTTGGCCTTGGAAACAGGCTGTCTC	13137
Qy	584	His 584	
Db	13138	CAC 13140	
RESULT 10			
AC084857			
LOCUS	AC084857	133683 bp	DNA linear HTG 22-NOV-2000
DEFINITION	Homo sapiens chromosome 11 clone CTD-3231NS map 11, WORKING DRAFT		
ACCESSION	AC084857.1	GI:11276215	
VERSION	AC084857.1	GI:11276215	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abram, H., Allen, N.,		
TITLE	Anderson, S., Barna, N., Bastien, V., Bede, F., Boguslavsky, L.,		
JOURNAL	Boukhalil, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,		
REFERENCE	Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,		
AUTHORS	Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Ferreira, P.,		

TITLE
JOURNAL
COMMENT

FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heald, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kanu, L., Karatas, A., Laroque, K.,
Lamatzes, R., Landers, T., Lechoczy, J., Levine, R., Lieu, C., Liu, G.,
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
McPheters, R., Melidrim, J., Menus, L., Mihova, T., Mlenaga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T.,
O'Donnell, P., O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, D.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L11038
Center clone name: 3231_N_5

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 126761 bases at least Q40
Consensus quality: 129952 bases at least Q30
Consensus quality: 131351 bases at least Q20
Insert size: 128000; agarose-fp
Insert size: 132483; sum-of-ctrls
Quality coverage: 6.1 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1459: contig of 1459 bp in length
* 1460 1559: gap of 100 bp
* 1560 2520: contig of 961 bp in length
* 2521 2620: gap of 100 bp
* 2621 6334: contig of 3714 bp in length
* 6335 6434: gap of 100 bp
* 6435 11913: contig of 5479 bp in length
* 11914 12013: gap of 100 bp
* 12014 20756: contig of 8743 bp in length
* 20757 20856: gap of 100 bp
* 20857 29735: contig of 8879 bp in length
* 29736 29835: gap of 100 bp
* 29836 38645: contig of 8810 bp in length
* 38646 38745: gap of 100 bp
* 38746 52193: contig of 13448 bp in length
* 52194 52293: gap of 100 bp
* 52294 65203: contig of 12916 bp in length
* 65210 65308: gap of 100 bp
* 65310 79796: contig of 14487 bp in length
* 79797 79896: gap of 100 bp
* 79897 95896: contig of 16000 bp in length
* 95897 95996: gap of 100 bp
* 95997 113413: contig of 17423 bp in length
* 113420 113519: gap of 100 bp

QY 361 GlySerThrpheasnlle----- 366
 Db 10258 GGCTCCACCTTCAACATCAGGTGTGCCAGAGCCAGGGGGCTGGTGGGAAGCCCTCT 10317
 QY 366 ----- 366
 Db 10318 AGTCAGGGGTCTGCCCTAGGAAGTAAATAGCACTAGTTAATGCATACAGTTGCTTCA 10377
 QY 366 ----- 366
 Db 10378 GTAAGTGCAGGCACTACTATGCTCTTATATAACATTAATATTTTTTCTCCCAAT 10437
 QY 366 ----- 366
 Db 10438 AATTCGTGTTTGTATCCCAAGTTTTCAGATAATTAAGTACAGTTTCAGAGAGTAAG 10497
 QY 366 ----- 366
 Db 10498 TTGTCCAGGCCACATAGCTACCAATGGTGCATTTGCTACTGAAAGACAGCCTGTGAT 10557
 QY 366 ----- 366
 Db 10558 CAGTGATGCATGGAAGCTTAGGACCTGGCTCTGTGCATCAGAACTATGTTTCTTTTC 10617
 QY 366 ----- 366
 Db 10618 TTTTGGACAGATATCTCCCTGTCTGCCAGGTGGAGCGAGTGGCGTGTGATCTTGCT 10677
 QY 366 ----- 366
 Db 10678 CACTGAACCTCCGGCTCTCGGGTTCAGATGATTCCTCTGCTTACGCTCCCAAGTACT 10737
 QY 366 ----- 366
 Db 10738 GGGATTACAGTGCCCAACACCAACTGGCTAATTTTGTACTTTAGATAGATAGAG 10797
 QY 366 ----- 366
 Db 10798 TTTCACATGTTGGCAGGCTGTCTCCAACTCTGACCAAGTAATCTGCCCTTTGGCC 10857
 QY 366 ----- 366
 Db 10858 TCCCAAAATGCTGAATATATAGGTGCAAAACTATGTTTTCGATTAAGCTAGATCTTG 10917
 QY 366 ----- 366
 Db 10918 GATGGAAATGGAAGTGGGGTTCCTGGGATGGGGAGGGGAGCAAAAGTCCAGCAGGC 10977
 QY 366 ----- 366
 Db 10978 AGCCAGGCATCACAGTAACCTCTGAATTGACTTTGTCTACCGAGTAAGGGGCTCAGG 11037
 QY 366 ----- 366
 Db 11038 CCACCACAGAGCCAGACTATATCCCAATGATGTCACATTCCTCGATTCCATCTGAATC 11097
 QY 366 ----- 366
 Db 11098 CCTCTTGAAGTGAAGTGGGTGAAGGGCTATCCAGCTGTCTCTTCTCCAGAGCAAC 11157
 QY 366 ----- 366
 Db 11158 AGAGTTGAAGTGCCTTGGAGAGTGTGGGCATGTCAAGGTTTCTCAAGGGTTTC 11217
 QY 366 ----- 366
 Db 11218 TTCCACGGTATCAGTGTCTGTCTGTCTTTCTTTTCTTTTCTTTTAAAGGAGT 11277
 QY 366 ----- 366
 Db 11278 TTCACTCTTGTGGCCAGAGCTGAGTGCAGTGCATATCTCGGCTCACTGCAACCTCC 11337

QY 366 ----- 366
 Db 11338 GCCTCCAGATTCAGCAATTCCTGCTCCAGCTCTGATAGTGGATTATAGTG 11397
 QY 366 ----- 366
 Db 11398 CCAGCCACCAAGCCCGCTAATTTTGTATTTTGTAGAGACAGTTTCACATGTTGGC 11457
 QY 366 ----- 366
 Db 11458 CAGCTGTCTGAAGTCTGACCTGACCTGATCCACCTCTCAGCTCCCAAGTCT 11517
 QY 366 ----- 366
 Db 11518 GGGATTACATGTTGAGCACCTGTGCTGCTGCTTGTCTTTTAAGAACCAATATCT 11577
 QY 366 ----- 366
 Db 11578 ACTAGCTCAATCGAGTTTAAGTACTATAGATAGTGAAGAGGTTGGGAAG 11637
 QY 366 ----- 366
 Db 11638 TGATCAATGAAGGCTGAGGCTTGCTTAGTCAAGAACTTTCTGAGAGTACTTGA 11697
 QY 366 ----- 366
 Db 11698 GCCCTACATGTCGTATACCCAGACAGCTGAAGTTGTTGAGGGTGGAGGGCTGAAA 11757
 QY 366 ----- 366
 Db 11758 CAGAACATTAAGCATAGACTTGTCTCCAGAGATGCACAATTTATGAGGAGCTCAA 11817
 QY 366 ----- 366
 Db 11818 ACCCAAGCTCAAACTCTGATACAAAGTACAAAGTATGATGCCAAAAAGGACAG 11877
 QY 366 ----- 366
 Db 11878 AACATGGAACAGATCATCTTGTCTGCTGGAGGGGCTTCAGCTGGTCTGAGCT 11937
 QY 366 ----- 366
 Db 11938 GAGCCATGGAATGGAAGAACTGAACTTGGGCAAGGCAAGCCATCTCTCTGTAG 11997
 QY 366 ----- 366
 Db 11998 ATAGCTTTCCTTCAGGGTAAGGTCTGGGGCTCCGGGATGCCTGTTGCTAGAGAT 12057
 QY 366 ----- 366
 Db 12058 CAAATTTCTTTGTGATGTCACCTCCAGTTGGAACCAAAATTCCTGCAATGCCAG 12117
 QY 366 ----- 366
 Db 12118 AGTCACTATGGGCTCATCTGAACCACTATGCAAGGCAAGGCTGTTTGTACTGCCT 12177
 QY 366 ----- 366
 Db 12178 GGAGTGAAGGGTTTACAGGGAAAGTAATGATGAGAGGCTTTACAGCCAGGGGG 12237
 QY 366 ----- 366
 Db 12238 TGGTTGCGGGGTTGATGTTAACTGTGTCAAGAGGAATCMACAAACGTAGAGTGAG 12297
 QY 366 ----- 366
 Db 12298 CTGGGCTGAGAGATCAAGGAGGTACAGAGCATCAGAGAGAGAGTGAAGCTGGG 12357
 QY 366 ----- 366
 Db 12358 GCATGTGAGGAAGAGGCTGTGCTTGGCTTGGCCAACTGAGAGAGAGGAGGGGGT 12417
 QY 366 ----- 366


```
QY 82 ----- 82
Db 5879 CGCCGGGCGCTGGCCCTGTGGCCGACGCTGGCCCTCGCCCTCGCCCGCC 5938
QY 83 -----GlyTherVal1ThrLeuVal1Val1ThrArgTrp1PheAsn1IleVal1
Db 5939 CCTCTCTGCCCGACGCTTCTACGTGACGCTGTGTGACCCGCTGTGTGAACAGTACGAG 5998
QY 99 AsnLeuPro1TrpProAspArgLeuMetSerLeuVal1SerGlyPheVal1GluGlyLysAsp 118
Db 5999 AACCTCCGCTGGCCCGACCGCCCTCATGAGCTGTGTGTGCGCTTGTGTGCAAGGCAAGGAC 6058
QY 119 GluGlnSerArgLeuLeuArgArgTrpLeuLeuArgTrpAlaAsnLeuGlyAsnVal1Leu 138
Db 6059 GAGCAAGGCGCGCTGCTGCGGCGCACGCTCATCGCTACGCGCAACCTGGGCAAGTGTCTC 6118
QY 139 IleLeuArgSerVal1SerThrAlaVal1IleVal1ArgPheProSerAlaGlnHisLeuVal1 158
Db 6119 ATCTGTGCGAGGCTCAGCACCGCAGTCTACAAAGCTTCCCGCAGCGCCGACGACCTGTGTG 6178
QY 159 Gln----- 159
Db 6179 CAAGCAGTGGGCGGACCGGAGCAACGGGAGGACCGCGGACAGGCGGCGGAGAT 6238
QY 159 ----- 159
Db 6239 GGGCGGCGAGAAATGAGATGGGTGAGCCAAAGTCCCGGACTCGGGGGATTGGGT 6298
QY 159 ----- 159
Db 6299 GAGCGCAGGAGTGGGGGTGTAGTCAAGATTTGGGGGTCCAAATTTGGCGGGACAGAGTGGG 6358
QY 159 ----- 159
Db 6359 TGTCTGAAGTGGGGCGAGGCCAGAGCCACCTCCAGAGCTAGAGAGTCTGAGCAGG 6418
QY 159 ----- 159
Db 6419 CTAAGGACCTTGAAGGATTAATGAAAGAGGTGACGCTTGGGAACTGTGAGTACT 6478
QY 159 ----- 159
Db 6479 AGGCTACTACTTCCCTCTGCCCTTGCCCTTGTGATCTCGGCTTTCACACTGTGAGAGTATG 6538
QY 159 ----- 159
Db 6539 GGACATTGGTCTGTGACACCCCTCAGCCTGGCTGACCTGTCTGTCTGTTAATAAGACAG 6598
QY 159 ----- 159
Db 6599 ACCCAGGCTAGGCGTGGTGGCTGTGCCTGTAAATCCAGTGTAGAGGCAAAAGTGTG 6658
QY 159 ----- 159
Db 6659 GAAGATGCTTGAAGCCAGCTGTTTGAAGCGCCCTGAGCAACATAGCGAGACCCCATTC 6718
QY 159 ----- 159
Db 6719 TCTACAAAACATTAAATAATAGCAGGCGCATGTGTGGGTGTGTGTGACAGTCTGAGGCTGA 6778
QY 159 ----- 159
Db 6779 GTATCGGAGGCTGAGGCGAGGAGATACATTGAGCCAGAGTTCAGGCTGACGTGCGC 6838
QY 159 ----- 159
Db 6839 TAAGATGACACCGCTGCACTCAACCTCGGTGACAGACGACGACCTTCTCTGGAATA 6898
QY 159 ----- 159
Db 6899 AATAAATACCTTCCACATGCTCAGCCAGAACAGCACTAGTAGTGTCTCAGAAATTT 6958

QY 159 ----- 159
Db 6959 TTTTGTGTGAAGAAAGAGATGGCAAGAGAGTGTGAGTTCCTATAGTACAGAG 7018
QY 160 -----AlaG1 161
Db 7019 TCGCGGCAATCCCTTCTGCAGGTTTCCACCCACCGCTTCTTCACTCCACTCTGCAGG 7078
QY 161 yPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsnMetPh 181
Db 7079 CTTTATGACTCCGGCAGAACACAGAGAGTGGAGAAACTGAGGCTTCCACACACACTGT 7138
QY 181 eTrpVal1ProTrpVal1TrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyGlyArg11 201
Db 7139 CTGGGTGCTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7198
QY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn----- 212
Db 7199 CCGGACCTATCTGTCTCTCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7258
QY 212 ----- 212
Db 7259 CCGCAGATGGGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7318
QY 212 ----- 212
Db 7319 GGGCCCTGAGGCTTCTCCAGAGCTGAGTGGGGTTGCAGAAATCTTTTCCACAGACAA 7378
QY 212 ----- 212
Db 7379 TCCACAGCCCGAGTGTCCCTTCCAGAGGCTTCCCTCTTCTTCCAAAGTGTGAGGT 7438
QY 212 ----- 212
Db 7439 CCGTGTCCCTTTGTATGATGAGAGAACGTGAGACACAAAGAGTTAGTGTGTGTGTGTGT 7498
QY 212 ----- 212
Db 7499 TGGCCACACACAGGAATGAGACCATAGGTACAGGCTTGTGTGTGTGTGTGTGTGTGTGT 7558
QY 212 ----- 212
Db 7559 GGGGAGCCGAGGTTGGGGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7618
QY 213 --GluMetAsn1ThrLeuArgTrpGlnGlySerGlyHisLeuTrpAlaTrpAspTrpLysSer1 232
Db 7619 AGGAGATGAAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7678
QY 232 LeProLeuVal1TrpThr----- 237
Db 7679 TCCCACTGTGTATATACACAGGTGAGACTAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 7738
QY 237 ----- 237
Db 7739 AGGCTAGAGAGACCAAGAAAGACAGCTGGGGTGGGAAAGGCTACCTAGAGGCTAAGTGC 7798
QY 237 ----- 237
Db 7799 TCCCTGTGGAGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7858
QY 237 ----- 237
Db 7859 AAGAGTCCAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7918
QY 237 ----- 237
Db 7919 AAGAGAGTGTGGCGGGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7978
QY 237 ----- 237
Db 7979 GGCAGGTGATACCTGAGGTGAGAGTTCGAGACCAAGCTGGGCAACATGTGTGAACCC 8038
QY 237 ----- 237
```


Db 153898 GAAACTGAGTCCAGAGAGAGAGATTCTCCATGCTACAGGCACATACAGGTCC 153839
QY 316 -----GlnValSerLeu 320
Db 153838 TCCCTGGGATGATCTTCTGTGGAGACTTCTGTCCCTGTGACAGAGGTGCCGTGTG 153779
QY 321 AlAlaLAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLys 340
Db 153778 GCTGTGATGATGATCACCAGAGACTGCTGCTCGATGGAGACCGGACATGTGATGGAATAG 153719
QY 341 ProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db 153718 CCGGAGCCACAGACCCCGCTACAGAGCTGCTCCGCCAGTTCCGTGAGCCCTCTTATG 153659
QY 361 GlySerTrpPheAsnIle----- 366
Db 153658 GGCTCACCTTCATCATCAGTGTGTGGCCAGAGCGGGGCTGGGGAGACCCCTCTCT 153599
QY 366 ----- 366
Db 153598 AGTCAGAGGGTCTGCTAGGAACCTAGAACTAGACTAGTTAATGCAATACAGGTTGCTTCA 153539
QY 366 ----- 366
Db 153538 GTAAGTGCAGGCACTGACTATGCTCTTTAATAACATTATTTTTTCCCTCCAT 153479
QY 366 ----- 366
Db 153478 AATTCTGTTTGTATCCCAAGTTTTCAGATAAATTAAAGTACAGGTTCCAGAGAGTAAAG 153419
QY 366 ----- 366
Db 153418 TTGTCCAAGCCACATAGCTACCAAAATGTGCTATTGCTACTCCAGAGACAGCTTGAT 153359
QY 366 ----- 366
Db 153358 CAGTGATGACAGTGAAGTTAGACCTGCTCTGTGATCCAGAACATATGTTTCTTTTC 153299
QY 366 ----- 366
Db 153298 TTTTTCAGACAGTATTCCTGCTCTGTCCGCCAGGTGGAGCGCAGTGCATCTTGAGCT 153239
QY 366 ----- 366
Db 153238 CACTGCAACCTCCGCTCCTGGGTTCAAGTATTCCTGCTTACGCTCCCCAGTACCT 153179
QY 366 ----- 366
Db 153178 GGGATTACAGGTGCCAACACACACACTGCTAATTTTGTACTTTTAGTAGAGATGAGG 153119
QY 366 ----- 366
Db 153118 TTTTCACCATGTGTGGCCAGGCTGGTCTCAACTCCTGACAGTAACTGCCCCGCTTGACC 153059
QY 366 ----- 366
Db 153058 TCCCAAAATGCTGGAATTATAGGTGCAAAACTATGTTTCTGATAGCTACGATGCTTG 152999
QY 366 ----- 366
Db 152998 GATGGAAATGGAATGGGGTTCCTGGGATGGGGAGGGGAGCAAAATGCCAGAGGC 152939
QY 366 ----- 366
Db 152938 AGCCAGGCCATCAGAGTACCTCTGAATGACTTGTCTTACCGAGTAAAGGGCTCAGG 152879
QY 366 ----- 366
Db 152878 CCACCCACAGACAGCACTTATCCCAATGTGCCACTTCCCTGATTCATCTGAATC 152819
QY 366 ----- 366
Db 152818 CCTCTGAGCTGCAGTGGGCTGAAGGCTATCCAGCTGTCTTCTCCCCAGAGCAAC 152759

QY 366 ----- 366
Db 152758 AGAGTTGAAAGTCCCTTGAGAGAGTGTGGGCACATGTACAGGTTCACTCAAGGGTTTC 152699
QY 366 ----- 366
Db 152698 TTCCAGGTATCCAGTGTGTTCTGCTTGTCTTCTTTTCTTTTAAAGGAGT 152639
QY 366 ----- 366
Db 152638 TTCACCTTGTGGCCAGAGCTGAGTGCAGTGGCATATCTCGGCTCACTGCACACTCC 152579
QY 366 ----- 366
Db 152578 GCTCCAGATTCAGCAATTCCTCTGCTCAGCTCCTGATGATGGATTATAGTG 152519
QY 366 ----- 366
Db 152518 CCAGCCACCAAGCCCGCTAATTTTGTATTTTAGTAGAGACAGTTCCACATGTGGC 152459
QY 366 ----- 366
Db 152458 CAGGCTGTCTGAACTCCTGACCTCAGGTATCCACCCTCCTCAGCCTCCAAAGTCT 152399
QY 366 ----- 366
Db 152398 GGGATTACATGTGTGAGCACTGTGCTGCTGCTTGTCTTTAAGAACCAATATCTCT 152339
QY 366 ----- 366
Db 152338 ACTAGACTGCATGAGTTAACTACAGCTATAGATAGTGTGAGAAAGGTGGGAAG 152279
QY 366 ----- 366
Db 152278 TGATCAATGAAAGCTGAGGCTTGCTTAGGTCAAAACATTCTCGAGATGACTTTGA 152219
QY 366 ----- 366
Db 152218 GCCCTACATGTGTGTACCCAGCAGCTGAAGTTGTTGAGGGGTGGGAGGCTGAAA 152159
QY 366 ----- 366
Db 152158 CAGAACGATTAAGCATGACCTGTCTCAAGGAATGCACAAATTTATGAGAGAGCTCAA 152099
QY 366 ----- 366
Db 152098 ACCCAAGCTCAAACTGTGATACAAAGTACAAAGTACTGATGTCCAGAAAAGGACAG 152039
QY 366 ----- 366
Db 152038 AACATGGAACACAGTATCTTGTGTGCTGCTGGAGGGGCTTCAGCTGGGTCTGAGCT 151979
QY 366 ----- 366
Db 151978 GAGCCATGGAACATGGGAAGATGTGAATCTTGAGCAAGGACAGCCATCTCTGTGTAG 151919
QY 366 ----- 366
Db 151918 ATTAAGCTTCCCTTGCAGGTTAAAGTCTGGGCTCCCGGATGCTGTGTGTAAGAGT 151859
QY 366 ----- 366
Db 151858 CAAATTTCTTGTGTGATGTACCTCCAGTTGGAACACAAATTCCTGGCATTTGCCAG 151799
QY 366 ----- 366
Db 151798 AGTCACTATGGGCTCACTGTAACCACTCATGACCAGGACCAAGTGTCTGACTGCTCT 151739
QY 366 ----- 366
Db 151738 GGAGTGAGGGGTTTTCAGGGGAAGTGAATGATGAGAGGCTTTTACAGCCAGGCGGGG 151679


```
OY 237 ----- 237
Db 156057 AAGATCCAGGCTCCTGCTGCGCCAGTCCAGTAGAGGCAATGTATTATCCCATATTA 155998
OY 237 ----- 237
Db 155997 AAGAGAGTTGGCGGGGTGAGTGGCTCATGCTGTATCCAGACACTTTGGAGCTGA 155938
OY 237 ----- 237
Db 155937 GCGAGGTGATCACCTAGGTGAGAGTGGAGACAGCCCTGGCCAAATGTTGAAACC 155878
OY 237 ----- 237
Db 155877 CATCTCTAGTAAAAATACAGATTAGCTGTGTGTGTGTGCACGCCCTGTAATCCAGCTAC 155818
OY 237 ----- 237
Db 155817 TTGGGAGGCTGAGGACAGAGAAATCGCTTGAACCCGGGAGGTGGAGGTTGCAGTAGCTGA 155758
OY 237 ----- 237
Db 155757 GATCATGCCACCTGCACTCCAGCTGGGCGCACAGCAAGACTGTGTCTCAACAAACAAA 155698
OY 237 ----- 237
Db 155697 CAACAAACAAACAAAGGGGTTAACAGAGCCCTAAGTCAATAGTGTGCAAGTCAGAA 155638
OY 237 ----- 237
Db 155637 CAAGGCTTGGTCTCCTGTCTAGACTCCAGCCCTGGAGCAATTCATTCAGGGTTC 155578
OY 238 ----- 245
Db 155577 CCACCTAGCCCTTGTGTACACATCTCTCTCTCTCCACAGGTGGTGTGCGGTGT 155518
OY 245 yfSerPheheLeuThrCysLeuValGlyArgGlnPheLeuAsnProAlaLysAlaTyrP 265
Db 155517 ACAGCTTCTCTCTGCTGTCTGTCTAGTTGGCGGCACTTCTGAACCCAGCCAAAGGCTTACC 155458
OY 265 toGlyHisGluLeuAspLeuValProValPheThrPheLeuGlnPhePheItyrV 285
Db 155457 CTGGCCATGAGCTGAGACCTCGTGTGTGCCGTCTTACAGCTTCTCGAGTTCCTTCTATG 155398
OY 285 aIGlyTyrLeu----- 288
Db 155397 TTGGCTGGCT-GAAGGTGGCCCTCTCAAGGGCCCTGCTGGGCTGAGGCAATGGCCAGAG 155339
OY 288 ----- 288
Db 155338 GGTCAITGCCACAGACTGCTGAGACGAGATGCAGTGTCAAGAAAGAGGTTCACGG 155279
OY 288 ----- 288
Db 155278 GTAGAAAGCAGCCAGCGGTGTGGCGCACACCTGTATCCAGCTACTCGGAGGCTGAG 155219
OY 288 ----- 288
Db 155218 GCAGAGAAATCGCTTGAACCCGGAGGCGGAGGTTGTGTGAGTTGAGATGTCGCACAG 155159
OY 288 ----- 288
Db 155158 CACTCAGCGCTGGGCAAAAGAAATGAATCTATCTCAAAAAACAACAACAACAACAACA 155099
OY 288 ----- 288
Db 155098 AAGCCCTAAGGTTCAAGAACCCCTGCCCTTAGAAGCAGATGGGCAACACCTCTCCTTA 155039
OY 288 ----- 288
Db 155038 TTCAAGATGCTGCTGTGGGCTGTCTTGTCTCACTCACTCAAGTGTGCTCCAGGATATCC 154979
OY 288 ----- 288

Db 154978 TCCACACACAGCAAACTACTCCGAACAGATGTTGAAATCACACAGTTTCTCTCACCTC 154919
OY 288 ----- 288
Db 154918 TTTATCTTCTCTCTCTCTGTGTGGCCACCACTCTCTCTCTCTCTCTCTCTCTCTCT 154859
OY 288 ----- 288
Db 154858 TTTTGGTAATGGGGGTGAAGTCTGTGTCTGTGTCTGTGTCTGTGTCTGTGTGTGTGT 154799
OY 288 ----- 288
Db 154798 ACACACACACACACACACACATACACACACACACACACACACATTCCTATTCCTTAAT 154739
OY 288 ----- 288
Db 154738 TCCCTCTGCCCCCCAGTTATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 154679
OY 288 ----- 288
Db 154678 CTTGAAATTTCTCCGGGGTCCCGCCAGTGGCTGCAGAGATGTCCTGAGCCCTTAAGCAG 154619
OY 288 ----- 288
Db 154618 ACGGTGTACCTCTTCGGGGCTTTGTAGGGCATTTTAGAGGTTCCTATCCAGAAATCT 154559
OY 288 ----- 288
Db 154558 GCCACCTAGACTGCCCTTTAGTTCAGCCAGCTTCAGTATATATCTGTTCATGAAT 154499
OY 288 ----- 288
Db 154498 GAATAAATTTATGCAACTCCAGTAAAGATACATGAGTGAGATAAAGCAGTACTCAGC 154439
OY 288 ----- 288
Db 154438 CGAGTGATACACTAGGACAGACACTGTGGGTGTTACAGGAAGACATGCTCAGACAGATTA 154379
OY 288 ----- 288
Db 154378 GAGGGGTGTGTCCAGAAAGTGTGTGGTCCACAAAGTGTGGGGCTGAGCCCTAAAC 154319
OY 288 ----- 288
Db 154318 TCTGCTTTGAAGACAGTGTGTCAGGACAGAAAGGGGTATGGGTGTGGAATACACACA 154259
OY 288 ----- 288
Db 154258 GCTGAGGTTTAAAGGGGAGACCTGCTTGAAGAGTTCTGCTGAGGGTTTACAGAGCT 154199
OY 289 ----- 305
Db 154198 CACCTGTCCCAAGGTGGAGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGAT 154139
OY 305 heGIuThrAsnTyrPrlLeuValAspArgAsnLeu----- 315
Db 154138 TTGAGACCACTGATGATGTGACAGGAATTTGCAGGTATGGGAGAGAGAAACCAT 154079
OY 315 ----- 315
Db 154078 ACCATGAGCTTCCCAAAATGGAGCCAAAGAGAGAGAGCCCACTGTCTGTAGGAGGCC 154019
OY 315 ----- 315
Db 154018 TCACAGTGAATGATCAACCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 153959
OY 315 ----- 315
Db 153958 ATTCTACCTCAATCTTTGAGGCTGACAGGACGACCACTCCCATTTTCACAGCAGG 153899
OY 315 ----- 315
```

Db 158277 TAGCAATGAAAAAATTTCTGAGGGAAGGCTGACATCATGTCCTCGAGCCCTG 158218
QY 82 ----- 82
Db 158217 CCGGGAGGGGAGGGGGCTGTGCGGATTTCTGGACACGAGGGGAGCCCGGGTGACA 158158
QY 82 ----- 82
Db 158157 GAACCTTGGGGCTGTGCGGCTCCATGCGAGGCTGTGCTGCTGCTGCTGCTGCGC 158098
QY 82 ----- 82
Db 158097 CTTCCAGAGAGGCTGGGGGCTAGGCCCGCTCGACAGAAAGCTGGAGAGCCGAGCAT 158038
QY 82 ----- 82
Db 158037 CCGCGGGGCTGGGGCTGGGCTGTGGCGGACGCTGGCCCTCGCCCTCGCCCGC 157978
QY 83 ----- 157978
GlyPheTyrValThrLeuValValThrArgTyrTrpAsnGlnTyrGlu 98
Db 157977 CCGCTCGCGCCAGGCTTCTAGGTGACGCTGGTCGTGACCCGCTGGTGAAACAGTAGAG 157918
QY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
Db 157917 AACCTGCGGTGGCCGACCGGCTCATGTAGCCGTGTGCGGCTGTGCGAAGGCAAGGAC 157858
QY 119 GluGlnSerArgLeuLeuArgArpThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
Db 157857 GAGCAAGGCGGGCTGTGCTGCGGCGCACGCTCATCCGCTAGCCCAACCTGGGCAACCTGTG 157798
QY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal 158
Db 157797 ATCTGTGCGCACGCTGACGACCGCAGTCTACAAGGCTTCCCGACGCGCCACACTGTG 157738
QY 159 Gln ----- 159
Db 157737 CAACGAGGTGGCGGACCGGAGCAACGGGAGCGGACGAGAGCCAGGCGCGAGAT 157678
QY 159 ----- 159
Db 157677 GGGCGGCGAGGAATGGAAGATGGGTGAGCCAAAGTCCCGCGGACACTCGGGGATTGGGT 157618
QY 159 ----- 159
Db 157617 GAGCCAGAGAGTGGGGTGTAGTCAAGATTGGGGGTCCAAATGGCGGAGACAGATCGGG 157558
QY 159 ----- 159
Db 157557 TGTCTGAAGGTGGGGGCGAGGCCAGAGCCACCTCCGAGAGTAGAGTCTGAGGACGG 157498
QY 159 ----- 159
Db 157497 CTAAGACCTTGAGGGATATGGAAGAAGGAGGAGCGGCTGGGAACCTGTGAGGATACT 157438
QY 159 ----- 159
Db 157437 AGGCTCTACTTCCCTCTGCCCCCTTCTGTATCTCCGTTTCCACTCTGAGAGTATG 157378
QY 159 ----- 159
Db 157377 GGAATTTGGTCTGTGACACCCCTCAGCCTGGGCTGACCTGTGCTGGTTAATAGACAG 157318
QY 159 ----- 159
Db 157317 ACCAGGCTAGGCGGTGGGTGTGCTGTATATCCAGTGTCTTAAAGAGGCAAGGTGG 157258
QY 159 ----- 159
Db 157257 GAAGATCGCTTGAGCCAGCTGTTTGAAGAGCCCTGAGACACATAGCGAGACCCCAATC 157198
QY 159 ----- 159
Db 157197 TCTACAAAAACATTAATAATTAGACAGGCAATGGTGGCGTGTGCTGACAGTCTGAGGCTGA 157138

QY 159 ----- 159
Db 157137 GTATGGGAGGCTGAGGAGAGAGATCACTTAGGCCAGACAGTCCAGGCTGAGTCCGC 157078
QY 159 ----- 159
Db 157077 TAAGATCGACCGCTGACATCCACCTCGGTGACAGAGCAGACCCCTTCTCTGGAAATA 157018
QY 159 ----- 159
Db 157017 AATAAATACCTGCCACATGCTCAGCCAGAACAGCAGTATAGTGTCTCAGAAATTT 156958
QY 159 ----- 159
Db 156957 TTTTGTGTGAAGAAGAGAGATGGCAAGAGAGTGTGAGTTCCTATAGTCAAGAG 156898
QY 160 ----- 156898
AlaGln 161
Db 156897 TGCCGGCCATCCCTTCTGACAGTTCTCCACCCAGCCGCTTCTTCACTCCACTCTGACAG 156838
QY 161 yPheMetThrProAlaGlnHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetPhe 181
Db 156837 CTTTATGACTCCGGCAGAACACAAAGCACTTGAGAAACTGAGCCTACACACAACTATGT 156778
QY 181 eTyrValProTyrPValTyrPheAlaAsnLeuSerMetLysAlaTyrPheGlyArgIle 201
Db 156777 CTGGGTGCCCTGGGTGTGTTTGTCCAACTGTCAATGAAGCGTGGGTGAGGTGCAAT 156718
QY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn ----- 212
Db 156717 CCGGAGCCCTATCTGCTCCAGAGCTGTGAAACGTAGCCCACTGTACAGACAGGCTG 156658
QY 212 ----- 212
Db 156657 CCGCAGATGGGAAGGCTGTGTCACAGGAACAAGTTTCTTACAAAGAGAAGCTT 156598
QY 212 ----- 212
Db 156597 GGGCCCTGAGGCTTTCGAGAGCCTGAGGTGGGTGCGAATCTTTCACAAAGCAA 156538
QY 212 ----- 212
Db 156537 TCCACAGCCGAGAGTGTCCCTTCTGAGAGCCCTCCCTTCTCTCAAGTCTGTAGGT 156478
QY 212 ----- 212
Db 156477 CTTGTTCCTTTTGTATAGTAGAAGTAGAGACACAAAGAGTTTATGAGCTTCCA 156418
QY 212 ----- 212
Db 156417 TGCCACACAGCCAGGAATGACCAATAGTACCAAGCCCTGTACTGAGAGAGGTGG 156358
QY 212 ----- 212
Db 156357 GGGCAGACCCAGGGTGGGGGAGAGTGTGTTGAAACCCCATCCCTCTTCTGCCCCC 156298
QY 213 --GluMetAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTyrIleSerI 232
Db 156297 AGGAGATGAACACCTTGCCTACTACAGTGTGACACCTGTATGCTCAGACTGATTAATA 156238
QY 232 leProLeuValTyrThr ----- 237
Db 156237 TCCACTGGTGTATACACAGGTGAGACTAGGCTGTGAGCTGCCCTTTTGGGAAACTG 156178
QY 237 ----- 237
Db 156177 AGGCTAGAGACCAAGAAAGCAAGTGGGTGGGAAGGCTCACTAGAGGCTAAGTGGC 156118
QY 237 ----- 237
Db 156117 TCCCTGGAGTTGGGTCCACACTTTGAAGTTGGGTCTGAGCTTGAAGTGCAGAGTTCT 156058

```

      *          17286 17385: gap of            100 bp
      *          17386   19170: contig of 1785 bp in length
      *           19171 19270: gap of             100 bp
      *           19271 28600: contig of 9350 bp in length
      *          28601 28700: gap of              100 bp
      *          28701 47671: contig of 18971 bp in length
      *         47672 47771: gap of               100 bp
      *        47772 105063: contig of 57292 bp in length
      *       105064 105163: gap of                100 bp
      *       105164 142181: contig of 37018 bp in length
      *      142182 142281: gap of                 100 bp
      *     142282 163915: contig of 21634 bp in length.
FEATURES
    source
        1..163915
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /chromosome="11"
           /map="11"
           /clone="RP11-810P12"
           /clone_id="RPCI-11 Human Male BAC"
misc_feature
    1..17285
       /note="assembly_fragment
       clone_end:SP6
       vector_side:left"
misc_feature
    19271..28600
       /note="assembly_fragment"
misc_feature
    28701..47671
       /note="assembly_fragment"
misc_feature
    47772..105063
       /note="assembly_fragment"
misc_feature
    105164..142181
       /note="assembly_fragment"
misc_feature
    142282..163915
       /note="assembly_fragment
       clone_end:T7
       vector_side:right"
BASE COUNT   43950 a   38090 c  37945 g   43329 t    601 others
ORIGIN
Alignment Scores:
Pred. No.:      8.27e-132                      Length:                   163915
Score:          1645.50                       Matches:                     528
Percent Similarity: 20.28%                    Conservative:                  0
Best Local Similarity: 20.28%                  Mismatches:                   6
Query Match:     52.74%                        Indels:                       2070
DB:              2                             Gaps:                         .7

US-09-622-964 -3 (1-585) x AC087451 (1-163915)
OY                               51 ArgLeuAlaLeuThrGluInglnglnLnuMetPhgLIuLyLeuThrLeuTyrcysASP 70
DDb |-----|
    158577 AGCGTGGCCCTCAGCAAAACAAGAAGCTTGATTTGAAGAAACTGTACTGTATTGGCAC 15851
OY |||||||||
    71 SerTYrlleGlNleuileerollerSerPheValLuu ----- 82
Dd |-----|
    158517 AGTCATCATCCAAGCTATATCCCCTTTCTTCCTGCtgcTGAgTTCCCCCTTCGcGTgt 158458
Oy |-----|
    82 ----- 82
Db 158457 CCggGTccCTGTGCCGCCCCCAAGGCTCcAgaCaGaCCAGCGAGGAGAtCaCaGAGAGAcTGGCg 158398
Oy |-----|
    82 ----- 82
Db 158397 GCaaGGGGcTtGGGaGGGGGGCGGGGGAaCGCCAGCGGAGGTGGCGcCTCTGTAgGG 158338
Oy |-----|
    82 ----- 82
Db 158337 AAAGGTcCGAGctcAGcCaNaTaGAaaActGAaGTAgtTAGacCTTAGagTAAGTAAGtcCTgccGT 158278
Oy |-----|
    82 ----- 82

```

```
QY 366 ----- 366
Db 189873 CAGGCCACCCACAGCAGCAGACTTATCCACATGTCCTCCACTGCTGATTCATCTG 189932
QY 366 ----- 366
Db 189933 AATCCCTCTGAGCTGCAGTGGGCTGAGGGCTATCCCACTGCTCTTCTCCAGGA 189992
QY 366 ----- 366
Db 189993 CAACAGATTGAAGTGCCTTGAGAGATTGGGACATGTCAGGGTTACTCAAGG 190052
QY 366 ----- 366
Db 190053 TTCTTCCACGGTATCCAGTGTCTGCTGCTGCTTCTTCTTTTCTTTTAAAG 190112
QY 366 ----- 366
Db 190113 GAGTTCACTCTGTTGCCAGAGCTGAGTGCAGTGGCATATCTGGCTCAGTCGAC 190172
QY 366 ----- 366
Db 190173 CTCCGCTCCAGATTCAAGCAATTCCTGCTCAGCCTCTGAGTACGTGGATTATA 190232
QY 366 ----- 366
Db 190233 GGTGCCAGCCACCAAGCCGGCTAATTTTGTATTTTGTAGTAGACAGTTCACCATGT 190292
QY 366 ----- 366
Db 190293 TGGCAGGCTGCTCGAATCTGACCTGACCTGATGATCCACCTCTCAGCCTCCCAAG 190352
QY 366 ----- 366
Db 190353 TCCTGGGATTACATGTGTAGCCACTGTGCTGCTGCTTCTTTAAGAACCAATA 190412
QY 366 ----- 366
Db 190413 TCCTACTAGACTGCAATGCAAGTTTAAGTACACTATAGATAGTGTGAGGAAGGCTTGGC 190472
QY 366 ----- 366
Db 190473 AAGGTGATCAATGAAGCTGAGGCTTGCTTAGGTGAGAAACATTTCTGAGAGTACT 190532
QY 366 ----- 366
Db 190533 TTGAGCCCTACATGCTGTACCCACAGACTGAAGGTTGTGAGGGGTGGGAGGCTG 190592
QY 366 ----- 366
Db 190593 AAAACAGAACGATAAGCATAGACCTGTCTCCAGAGAAATGACAAATTTATGAGGAGAC 190652
QY 366 ----- 366
Db 190653 TCAAAACCAAGTCTCAAACTGTGATACAGGTACAAAGTACTGATGTCCAGAAAGG 190712
QY 366 ----- 366
Db 190713 ACAGAACATGAGACACATCTTCTTGTCTGCTGGAGGCGGCTTCCAGCTGGGCTG 190772
QY 366 ----- 366
Db 190773 AGCTGAGCATGAGAACATGGAAGATCTGAACCTTGGCAAGGCGAGGCATACTCTCTG 190832
QY 366 ----- 366
Db 190833 GTAGATAAGCTTTCCCTTGGCAGGGTAAAGTCTGGGGCTCCGGGAGTCCGTGTGAG 190892
QY 366 ----- 366
Db 190893 AAGTCAATTTCTCTTTGTGATGTCACTCCAGTGTGAACCAAAATTCCTGGCATTC 190952

QY 366 ----- 366
Db 190953 CCAGAGTCACTATGGGCTCATCTGAACACTCATGCCCAGGGCACCAGTGTTCGACT 191012
QY 366 ----- 366
Db 191013 GCCTGAGTAGAGGGCTTTTACAGGGGAAGTGAATGATGAGAGGCTTTACAGCCAGC 191072
QY 366 ----- 366
Db 191073 GGGGTGTTCCGGGGGTTGGATGTTAATCTGTGTCAAGAGGCAATCAACAACAGTGAG 191132
QY 366 ----- 366
Db 191133 TGAGTGGGCTGAGAGGATCAACCGGAGGTACAGGACAGATCAGAGAGAGGTGAGAG 191192
QY 366 ----- 366
Db 191193 TGGGCGATGAGGAGAACGGTGTGGCTTGGCTTGGCCAACTGAGAGAGAGCGG 191252
QY 366 ----- 366
Db 191253 GGGTAAGGAGAGTAAGGCCAGTGTGTGCTTGTCCACTGCTCAGCCCTGCATCT 191312
QY 367 ----- 367
Db 191313 CCTGTTCTTTCCAGCTGACAAAGAGAGATGAGTTCAGGCCCAATCAGAGAGAGCA 191372
QY 382 ----- 382
Db 191373 GAGGATGCTCAGCGTGCATATTGGCCGCTTCCAGGCTGCGACGTCCATGATATCA 191432
QY 402 ----- 402
Db 191433 TCCTCCAGGCAAACTCAAGSACCAACTACTGTGGCCCAAGAGGAATCCCTCTCCA 191492
QY 422 ----- 422
Db 191493 GAGGGCCGCCCCAAAACCAACAGGCAAGCCAAACGTTAGGGGCCAGAGACAA 191552
QY 442 ----- 442
Db 191553 CAAGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGTATCAAGGCC 191612
QY 462 ----- 462
Db 191613 AGGCTACTACAGTGGCCACAGACGCCCTCAGCCCACTCCATGTCTTCCCTAGA 191672
QY 482 ----- 482
Db 191673 ACCATCAGCCCGTCACAAAGCTTCACAGTGTCAAGGCATAGACCAAAAGCAAGCTT 191732
QY 502 ----- 502
Db 191733 AAAGACTGTGAGTCTGGGGCCAAAGAAAGTTTGAATTTGCTCTCAGAGAGGAGGGCC 191792
QY 522 ----- 522
Db 191793 CTTGATGAGACCCAGAAAGATCTCAAGTAGGAGGAAAACTGTGGAGTTAACTGAC 191852
QY 542 ----- 542
Db 191853 GATATGCGCAGACATCCCGGAAATCACTCAAGAACCTTTGGACAAATACCAACCA 191912
QY 562 ----- 562
Db 191913 CATACACTACACTCAAAAGATCATAGGATCTTATGCGCTTGGAACAGAGTCTGT 191972
QY 582 ----- 582
Db 191973 CCTCCAC 191979

RESULT 8
```

Db 187653 AACAAACAAAGCCCTTAAGGTTGAGAGGCCCTTGCCCTTTAGAGGACAGAGGGGCAC 187712
OY 288 ----- 288
Db 187713 CACCTCTCCTTATTCAGATGCGCTGTGGCTGTCTTGTCTTCTCACTCAAGTGGCTTGT 187772
OY 288 ----- 288
Db 187773 CCAGGTATTCCTCCACACAGCCAACTACTCCGAACAGATGTTCTGAATCAGACAGTTT 187832
OY 288 ----- 288
Db 187833 TCTGTCCACCTTATCTTATCTTCTTCTTCTGTGGCCACCACCTCTCTCCTTCCCTA 187892
OY 288 ----- 288
Db 187893 CCTTCCCTTATTTTGTGTAATGGGGGTGTAAAGTCTGTCTGTCTGCTTCTGTCTCACTG 187952
OY 288 ----- 288
Db 187953 TGACACACACACACACACACACACACACACACACACACATTCCTATTCCTCT 188012
OY 288 ----- 288
Db 188013 AAATTCCTCCCTGCCCCCAGTTATCTTGTGTTCTGAGATCAAAACAAATCAGACTTT 188072
OY 288 ----- 288
Db 188073 TATGCTGAAATTCCTCAGAGGGTCCCCAGTGGCTGCAAGATTCCTCCCTGGACCCCTAAG 188132
OY 288 ----- 288
Db 188133 GCAGAGCGGTGCTACCTCTTCGGGGCTTTGTTAGGGCAATTGAGGTGCTATCCAGGA 188192
OY 288 ----- 288
Db 188193 AATGCCCACTAGACTGCCCTTTAGTTCAGCCCACTTCAGTAAATATCTCTGTTGAT 188252
OY 288 ----- 288
Db 188253 GAATGAATAAATTAATGCAACTCCAGTAAATATCATGAGTGAATAAAGCAGTACT 188312
OY 288 ----- 288
Db 188313 CAGCGAGTGAATACACTAGAGCAGAGCTGTGGTGTTCAGGGAGGAGTGGCTCAGAGA 188372
OY 288 ----- 288
Db 188373 GTTAGAGGGGCTGTGTCCAGAAAGTGTGGTGGCCCAAGTGTGGGGGCTGAGCCCT 188432
OY 288 ----- 288
Db 188433 AAATCTGCTTTTGAAGAGAGTGTGTCAGCAGAAAGGCTTCATGGGGTGTGGAATAGC 188492
OY 288 ----- 288
Db 188493 AGCAGTGAAGTTAAAGGGGGAAGCTGGCTTTGAGAGTTCCTCGAGGGTTTACAGA 188552
OY 289 ----- 188552
Db 188553 GCTTCACCTGTCCCAAGGTGGCAGAGAGCTCATCAACCCCTTTGGAGAGGATGATGAT 188612
OY 304 AspPheGluThrAsnTPrIleValAspArgAsnLeu----- 315
Db 188613 GATTTTAAAGCACTGATGATTGTGACAGAAATTTGCAAGGTATGGGAGAGAGAGAAA 188672
OY 315 ----- 315
Db 188673 CCATACCATGAGCTTCCCAAGTGAACCCAAAGAGAGAGACCCACTGTTCTGTAGGA 188732
OY 315 ----- 315

Db 188733 GGCTCAGAGTGAATGATCAACCTTCCCTCTCTCTCTGACAGCAGTCAATTCATCA 188792
OY 315 ----- 315
Db 188793 CAGATTCACCTCAATCTTTGAGGCTGCAGGAGGACCAACCAATTCCTCCATTTCCACAG 188852
OY 315 ----- 315
Db 188853 CAGGAAACTGAGGTCCAGAGAGAGAGAGATTCCTCAAGTCAATCAGCACAATACAG 188912
OY 316 ----- 316
Db 188913 GTCTGCTGGATGATCTTCTGTGGACTTCTTCTGCTCCCTGTGACCAAGTGTCCCT 188972
OY 319 ULeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetIleTyrTrpAs 339
Db 188973 GTGGGCTGTGATGATGATGACCAAGACCTGCTCGATGAGCGGACATGTACTGGAA 189032
OY 339 NlysProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerP 359
Db 189033 TAAAGCCGAGCAGAGCCGCCCTTACACAGCTCTCCGCCAGTTCGCTGACCTCTCT 189092
OY 359 eMetGlySerThrPheAsnIle----- 366
Db 189093 TATGGGCTCCACTTCAACATTCAGTGTGGCCAGAGCCAGAGGGGCTGGTGGAGACCCC 189152
OY 366 ----- 366
Db 189153 TCCTAGTGAAGGGTCTGCTGAGACTTGAATACACTAGTTAATGATACAGGTGC 189212
OY 366 ----- 366
Db 189213 TTCAGTAACTGACAGCACTGACTATGCTCTTATTAACATTAACTATTTTTCCTCC 189272
OY 366 ----- 366
Db 189273 CAATAATCTGTGTTTGTATCCCAAGTTTTCAGATTAATTAAGTACAGGTTCAGAGAG 189332
OY 366 ----- 366
Db 189333 TAAATTGTCAAGGCCACATAGCTTACCAATAGTGCATTTGCTACTCGAGACAGCCTA 189392
OY 366 ----- 366
Db 189393 TGATCAGTGAAGTGAAGCTTAGACCTGTGCTCTGTCAATCAGAACTATGTTTCT 189452
OY 366 ----- 366
Db 189453 TTTCTTTTGAAGAGATATCTCGCTGTGTGCCCAAGTTGAGCGCAGTGGCTGATCTT 189512
OY 366 ----- 366
Db 189513 GGTCTACTCAACCTCGCTCTGTGGTTCMAAGTATCTCTGCTTCAGAGCTCCCACT 189572
OY 366 ----- 366
Db 189573 AGCTGGATTAACAGTCCCAACAACAACACTGCTAATTTTGTACTTTTATAGAGAT 189632
OY 366 ----- 366
Db 189633 GAGGTTTCAACATGTTGGCAGAGGTGTCTCAAACTCCTCAACCAATATCTGCCCGCTTT 189692
OY 366 ----- 366
Db 189693 GGCCTCCCAAAATGTGAATTAATAGTGTCAAAACTATGTTTCTGATTAAGCTACAGAT 189752
OY 366 ----- 366
Db 189753 CTTGATGGGAAGTGAATGGGGTTCCTGGGATGGGGAGGGGAGCAAAAGTCCACG 189812
OY 366 ----- 366
Db 189813 AGCAGCCAGGCCATCAGAGTACCTCTGAAATTAATGTTGTGCTTACGAGTAAGGGCT 189872

[illegible]

```
repeat_region      /rpt_family="Alu"
                    82013..82288
                    /rpt_family="Alu"
                    complement(83419)..83600)
repeat_region      /rpt_family="MIR"
                    complement(87320)..87624)
repeat_region      /rpt_family="Alu"
                    complement(87710)..87987)
repeat_region      /rpt_family="Alu"
                    complement(88186)..88472)
repeat_region      /rpt_family="Alu"
                    88822..89049
                    /rpt_family="Alu"
                    89224..89308
                    /rpt_family="MER20"
                    89729..90487
                    /rpt_family="Alu"
                    90504..90618
                    /rpt_family="Alu"
                    complement(93313)..93952)
repeat_region      /rpt_family="Alu"
                    96938..97517
                    /rpt_family="Alu"
                    complement(10419)..104514)
repeat_region      /rpt_family="Alu"
                    complement(109738)..109979)
repeat_region      /rpt_family="Alu"
                    110022..110141
                    /rpt_family="Alu"
                    complement(111091)..112016)
repeat_region      /rpt_family="Alu"
                    113210..113419
                    /rpt_family="MIR"
                    complement(114951)..115005)
repeat_region      /rpt_family="MIR"
                    115685..115791
                    /rpt_family="MIR"
                    116870..117005
                    /rpt_family="MIR"
                    complement(117877)..118172)
repeat_region      /rpt_family="Alu"
                    118352..118600
                    /rpt_family="Alu"
                    118679..118813
                    /rpt_family="Alu"
                    complement(121233)..121520)
repeat_region      /rpt_family="Alu"
                    122668..123071
```

Alignment Scores:

```
Pred. No.: 2.55e-132      Length: 196080
Score: 1652.50      Matches: 529
Percent Similarity: 20.31%      Conservative: 0
Best Local Similarity: 20.31%      Mismatches: 5
Query Match: 52.96%      Indels: 2071
DB: 9      Gaps: 7
```

US-09-622-964-3 (1-585) x AC004228 (1-196080)

```
QY 51 ArgLeuAlaLeuThrgInglInglInLeuMerPhelInLysLeuThrlLeuTyCysasp 70
Db 184173 AGGGTGGCCCTCACGGAGAACACACAGCTGATGTTGAGAAACAGACTGTATTCGCAC 184232
QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu-----82
Db 184233 AGCTACATCCAGCTCATCCCATTTCTTGTGTGGGTGAGTTCCTCCCTTCGCGCTGT 184292
QY 82 -----82
Db 184293 CCGGGTCCTGTGGCCGCCAGGCTCCAGACAGGCCAGGGAGAGATCAGAGAGAGTGG 184352
QY 82 -----82
```

```
Db 184353 GCAGAGGGCTGGGGAGGGGGGGGGAACCCAGCGGAGGTGGCGCTCTGTGAGG 184412
QY 82 -----82
Db 184413 AAAGGTGGGACTGCAGACCCAGAGAACTGAATTAGCGTTAGAGTTCCTGCCGT 184472
QY 82 -----82
Db 184473 TAGCAATGAAGAACCCCATTTTCTGAGGGAAGCCCTGACATCATGTCTCCCTGAGCCCTG 184532
QY 82 -----82
Db 184533 CCGGGAGGGAGAGGGGTCTGGCGGATTTCTGGGACACAGAGGGAGCCCGGGTGACA 184592
QY 82 -----82
Db 184593 GAACCTTGGGGCTCTCGCGCTCCATGCGAGGCTCTGCTGCTCTGCTCCGAGCGC 184652
QY 82 -----82
Db 184653 CTTCAGAGAGGGCTGGGGGCTAGGCCCGCTCGACAGAGAAAGCTGAGAGCCGAGCAT 184712
QY 82 -----82
Db 184713 CCGCGGGGCTGGGCCCTGGGCTTGCCGCGAGCTGGCCCTCGCCCTCGCCCGCCCGC 184772
QY 83 -----GlyPheTyrValThrLeuValValThrArgTyrTrpAsnGlnTyrGlu 98
Db 184773 CCTCTCGCCCGGCTTCTACTGTACGCTGCTGCTGACCCGCTGGGGAAACAGTACGG 184832
QY 99 AsnLeuProTyrTrpProAspArgPheLeuMerSerLeuValSerGlyPheValGlnTyrAsp 118
Db 184833 AACCTGGCGTGGCCCGACCGCTCATGAGCGTGGTGGGCTGTGTGGAAGCAGAGAC 184892
QY 119 GluInSerArgLeuLeuArgArgPheLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
Db 184893 GAGCAAGCCCGGCTGCTGGGGGACGCTCATCCGCTACGCCAACCTGGGCAACCTGCTC 184952
QY 139 IleLeuArgSerValSerThrAlaValTyrIleAspPheProSerAlaGlnIleLeuVal 158
Db 184953 ATCTGGCGCAGCGTCAGACCGCAGTGTCAAGAGCCCTTCCCGAGGCCACCTGCTG 185012
QY 159 Gln-----159
Db 185013 CAACGAGGTGGCGGACCGGAGCAGGAGCAGGAGCAGGAGCCAGAGCCAGGCGCGAGAT 185072
QY 159 -----159
Db 185073 GGGCGCGCAGGAATGGAAGATGGGTGAGCCAAAGTCCCGGACTCGGGGATTTGGGT 185132
QY 159 -----159
Db 185133 GAGACGAGAGTGGGGTGTGTCAAGATTTGGGGGTCCAAATTTGGCGGGAGACAGTGGG 185192
QY 159 -----159
Db 185193 TGCTGAAGTGGGGCGAGGCCAGACCCACCTCCGAGAGTGAAGTGTGAGGCAAGG 185252
QY 159 -----159
Db 185253 CTAAGGACCTTGAGGATTAATGGAAGAAAGGCTGAGCGCTTGGAACACTGGTGAAGTACT 185312
QY 159 -----159
Db 185313 AGGGTCTACTTCCCTGTGCGCTTGCCTCTTGTATCTCCGTTTCCACTCTGGAAGTATG 185372
QY 159 -----159
Db 185373 GGAATTTGGTCTCTGACACCCCTCAGCCTGAGCTGACCTGTCGTGTTAATTAAGACAG 185432
QY 159 -----159
Db 185433 ACCCAGGCTAGGGGTGTGGTCTGCTGCTGTATATCCAGTGTCTTTAGGAGGCAAGGTGG 185492
```

SOURCE	Homo sapiens.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1	Evans, G.A., Athanasiou, M., Agunayo, P., Armstrong, D., Basil, M., Butler, J., Bunesster, R., Card, P., desaliboat, F., Dunn, J., English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T. HTGS Submissions
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 196080)
AUTHORS	Evans, G.A., Athanasiou, M., Basil, M., Bradbury, P., Brignac, S., Bunesster, R., Davis, C., English, C., Franklin, T.L., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Narayanaswamy, V., Newton, J., O'Brien, K., Patel, P., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T. and Wilson, R.
JOURNAL	Direct Submission
TITLE	Submitted (26-FEB-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
REFERENCE	3 (bases 1 to 196080)
AUTHORS	Evans, G.A., Athanasiou, M., Agunayo, P., Armstrong, D., Basil, M., Butler, J., Butler, C., Card, P., desaliboat, F., Dunn, J., English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Waller, K. and Ward, T. Direct Submission
JOURNAL	Submitted (24-FEB-1999) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
COMMENT	On Feb 24, 1999 this sequence version replaced gi:2911733. IMPORANT: This submission contains the entire insert of clone PDJ519013. PDJ519013 comes from the RPO1-3 PAC library constructed at the Roswell Park Cancer Institute by the Pletier de Jong group. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST ANNAK. This region spans over 1.5 Mbp. MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA), STS D11S699 and WI-7524
FEATURES	MAPPED CLONE OVERLAP: PACs PDJ466a11 and PDJ756b9. Location/Qualifiers
source	1. 196080
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
repeat_region	complement(2210..2482)
	/rpt_family="Alu"
repeat_region	3457..3743
	/rpt_family="Alu"
repeat_region	complement(20145..20738)
	/rpt_family="Alu"
repeat_region	complement(23194..23599)
	/rpt_family="Alu"
repeat_region	complement(27367..27685)
	/rpt_family="Alu"
repeat_region	28047..28457
	/rpt_family="Alu"
repeat_region	complement(29390..29565)
	/rpt_family="Alu"
repeat_region	complement(31076..31255)
	/rpt_family="Alu"
repeat_region	complement(34948..35045)
	/rpt_family="Alu"
repeat_region	35232..35538
	/rpt_family="Alu"
repeat_region	complement(36242..36540)
	/rpt_family="Alu"
repeat_region	37238..37322
	/rpt_family="Alu"

repeat_region	/rpt_family="Alu"	complement(38555	.38650
repeat_region	/rpt_family="Alu"	complement(40483	.41164
repeat_region	/rpt_family="Alu"	42042	.42016
repeat_region	/rpt_family="MIR"	42677	.42817
repeat_region	/rpt_family="LTR12"	42856	.43157
repeat_region	/rpt_family="Alu"	43148	.43380
repeat_region	/rpt_family="LTR12"	complement(43734	.43999
repeat_region	/rpt_family="LTR12"	complement(46989	.47169
repeat_region	/rpt_family="MIR"	complement(48046	.48346
repeat_region	/rpt_family="Alu"	48482	.48737
repeat_region	/rpt_family="Alu"	complement(48893	.49037
repeat_region	/rpt_family="L1"	49118	.49292
repeat_region	/rpt_family="Alu"	49791	.50059
repeat_region	/rpt_family="Alu"	50436	.50501
repeat_region	/rpt_family="Alu"	complement(51147	.51441
repeat_region	/rpt_family="Alu"	complement(52700	.52979
repeat_region	/rpt_family="Alu"	53281	.53568
repeat_region	/rpt_family="Alu"	54071	.54204
repeat_region	/rpt_family="Alu"	54306	.54793
repeat_region	/rpt_family="Alu"	54807	.55061
repeat_region	/rpt_family="Alu"	complement(64533	.64826
repeat_region	/rpt_family="Alu"	complement(65245	.65783
repeat_region	/rpt_family="Alu"	complement(68465	.68732
repeat_region	/rpt_family="Alu"	complement(70191	.70476
repeat_region	/rpt_family="Alu"	70560	.70692
repeat_region	/rpt_family="Alu"	72294	.72470
repeat_region	/rpt_family="Alu"	complement(72800	.73095
repeat_region	/rpt_family="Alu"	73424	.73656
repeat_region	/rpt_family="Alu"	73784	.74071
repeat_region	/rpt_family="Alu"	74652	.74935
repeat_region	/rpt_family="Alu"	complement(75265	.75514
repeat_region	/rpt_family="Alu"	complement(76086	.76353
repeat_region	/rpt_family="Alu"	complement(78038	.78312
repeat_region	/rpt_family="Alu"	78513	.78734
repeat_region	/rpt_family="Alu"	78866	.79180
repeat_region	/rpt_family="Alu"	complement(79384	.79673
repeat_region	/rpt_family="Alu"	complement(81338	.81690)


```
QY 288 ----- 288
Db 29033 CCAGTATTCCCTCCACACACCAACCACTCCGAACAGATGTTGTAATCACACAGTTT 28974
QY 288 ----- 288
Db 28973 TCTCTCAACCTCTTATCTTTCCTTCTCTGTGCCACCCACTCTCTCTCCCTTCTA 28914
QY 288 ----- 288
Db 28913 CCTTCCTTATTTTTTGTGAATGGGGGTGAAGTCTCTGTCTCTGCCCCCTCTGTCACTG 28854
QY 288 ----- 288
Db 28853 TGACACACACACACACACACACACACACACACACACACATCTTATTCCTCT 28794
QY 288 ----- 288
Db 28793 AAATTCCTCCCTGCCCCCAGTTATCTTGTGTTCTGCAGATCAAAACAAATCACACTT 28734
QY 288 ----- 288
Db 28733 TATGCTGAATTCCTCCAGGGTGCCCGCAGTGCCCTGAAGATGTCCTCGACCCCTAAG 28674
QY 288 ----- 288
Db 28673 GCAGACGGGTGTCACTCTTGGGGCTTTGTTAGGGCAATTTAGAGGTGCTATCCAGA 28614
QY 288 ----- 288
Db 28613 ATGCCCCACCTAGACTGCTTGTAGTTACGCCACAGTTTCAATATATCTCTGTGAT 28554
QY 288 ----- 288
Db 28553 GAATGAATAAATTTATGCACTCCAGTACAGATAGAGGTGAGATPAAAGCAGTGAAT 28494
QY 288 ----- 288
Db 28493 CAGCCGAGTATACACTCAGGACAGAGCTGTGGGTTCAGGGAAGAGACTGGCTCAGAGA 28434
QY 288 ----- 288
Db 28433 GTTAGAGGGCTGTGTCAGAAAGTGTGGTGCCCAAGTGTGGGGGCTGAGCCCT 28374
QY 288 ----- 288
Db 28373 AAACCTGCTTGAAGACAGTGTGTCAGGACAGAGGGCTTCATGGGGTGTGAATATGC 28314
QY 288 ----- 288
Db 28313 AGCAGTGAAGTTTAAAGGGGAAGCTGGCTTGAAGAGTTCTGCTGAGGGTTTACAGA 28254
QY 289 ----- 289
Db 28253 GCCTCACCTGTGCTCCCAAGGTGGCAGAGCAGCTCATCAACCTTTGAGAGATGATGAT 28194
QY 304 AspPheGluThrAsnTPIleValAspArgAsnLeu----- 315
Db 28193 GATTTTGAAGACCAACTGGATTTGTCAGACGAAATTTGAGGTATGGGAGAGAGAGAGAAA 28134
QY 315 ----- 315
Db 28133 CCATACCATGAGACCTTCCCAAGTGAGACCAAAAGAGAGACCCCACTGTCTGTAGGA 28074
QY 315 ----- 315
Db 28073 GGCCTCACAGTGAATGATCAACCTTCCCTCTCTCCCTCGAGCAGTCAATCACTCA 28014
QY 315 ----- 315
Db 28013 CAGGATTCACACTCAATCTTTGAGGCTGAGGACAGGACCACTCCCATTTTCACAGG 27954
QY 315 ----- 315

Db 27953 CAGGAAACTGAGTCCAGAGAGAGGAGAGATTCCTCAAGTCAATCAGCACATCAAG 27894
QY 316 ----- 319
Db 27893 GTCTGCTGGGATGATCTTCTGTGGACCTTCTGTGCTCCCTGGTGACAGGTGCTCT 27834
QY 319 UleuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAs 339
Db 27833 GTTGGCTGTGATGATGATGCACACAGACCTGCTGATGAGAGCCGACATGTACTGGAA 27774
QY 339 nLysProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgAlaSerPhe 359
Db 27773 TAAAGCCAGCCACAGCCCCCTTACACAGCTCTCCGCCAGTTCCGTGAGCCTCTT 27714
QY 359 eMetGlySerThrPheAsnIle----- 366
Db 27713 TATGGGCTCACCTTCAACATCAGGTGTGCCAGAGCCAGGGGCTGGGTGGAAAGCCC 27654
QY 366 ----- 366
Db 27653 TCCTAGTGCAGGGGTCTGCTAGGAATTAAGATAGCACTAGTAAATGATACAGGTTGC 27594
QY 366 ----- 366
Db 27593 TTCAGTAAGTGCAGGACACTGATGCTCTTATATAACATTAATTTTTCCTCC 27534
QY 366 ----- 366
Db 27533 CATTAATTCGTGTTGTTATCCCAAGTTTTCAGATTAATTAAGTACAGGTTACAGAGAG 27474
QY 366 ----- 366
Db 27473 TAACTGTCCAAGGCCACATAGTACCAAAATGTGCAATTTGCTACTGCAAGACAGCTA 27414
QY 366 ----- 366
Db 27413 TGATCAGTATGATGATGAGGAAAGTTAGAGACCTGTTCATCCAGAACTATGTCTCT 27354
QY 366 ----- 366
Db 27353 TTTCTTTTGAGACAGTATCTGCTGTGTGCCAGGTTGAGAGCCAGTGGCTGATCTT 27294
QY 366 ----- 366
Db 27293 GGCTCACTGCAACCTCCGCTCTGCGTTCAAGTATCTCTGCTTCAAGCTCCAGCT 27234
QY 366 ----- 366
Db 27233 AGCTGGATTAACAGGTGCCACACACAACTGCTAATTTTGTACTTTTATAGATAGAT 27174
QY 366 ----- 366
Db 27173 GAGTTTCACACATGTTGCCAGAGCTGTGCTCAACTCTGACCAAGTATCTGCCGCTTT 27114
QY 366 ----- 366
Db 27113 GGCTCCCAAAATGTGGAATTAAGTGTCAAAACTATGTTTCTGTATAAGTACGATG 27054
QY 366 ----- 366
Db 27053 CTGTGATGGAAAGTGAAGTGGGGTTCCTGGATGGGAGGAGGCGACAAAGTCCAGC 26994
QY 366 ----- 366
Db 26993 AGGCACACAGGCATCAGAGTACTCTGTAATTTGACTTTGTCTTACCGAGTAAGGGCT 26934
QY 366 ----- 366
Db 26933 CAGGCCACACAGCAGCACTATTCACCATATGATGTCCTCCATGATTCATCTG 26874
QY 366 ----- 366
```

Db 31253 TCCTCAAAAACATTAAAAATTAGCAGGCGCATGCTGGCGTGTGCTGTAGTGTGAGGCTGA 31194
QY 159 ----- 159
Db 31193 GTATCGGGAGGCTGAGCGAGAGGATCATCTTGAGCCAGCAGTTCCAGGCTGCATGCGC 31134
QY 159 ----- 159
Db 31133 TAAGATGCGACCGCTGCATCTCAACCTCGGTGACAGAGCCAGACCCTTTCTTGGAATA 31074
QY 159 ----- 159
Db 31073 AATAATACCTCCCATCATCTCATGCCAGAACAGCAGCTAGAGTGCTCAGAAATT 31014
QY 159 ----- 159
Db 31013 TTTTGTGTGAAGAAAGAGATGGCAAGAGTGTGAGTTCTTAATAGTCAGCAGC 30954
QY 160 ----- Ala1 161
Db 30953 TGGCGGCGATCCCTCTGCAAGTTCTCCACCCAGCGCTTCTCACTCCACTCTGCAGC 30894
QY 161 yPheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisMetPh 181
Db 30893 CTTTATGACTCGGCGAAGACACAAAGCAGTTGGAAACTGAGCCCTACACACAAACATGTT 30834
QY 181 eTrpValProTyrValTyrPheAlaMetLeuSerMetLysAlaTyrLeuGlyClyArg11 201
Db 30833 CTGGGTGCGCTGGGTGGTGGTTGGTCCCACTGTCAATGAAGCGCTGGCTGGAGTGAAT 30774
QY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn ----- 212
Db 30773 CCGGACCTTATCTGCTCTCAGAGCCTGTAACGTGAGCCCACTGTACAGACAGGCTG 30714
QY 212 ----- 212
Db 30713 CCGCAGAGTGGGAAAGGTTGTGTCCACAGAAAGAGTTCTTCAAAAGAGAGCCTT 30654
QY 212 ----- 212
Db 30653 GGGCCCTGAGGGGTCTCCGAGAGCGGAGTGGGGTTGCAAATCTTTTCCAAACAGCA 30594
QY 212 ----- 212
Db 30593 TCCACAGCCGAGGTGCTCCCTTCTCAGAGCCCTCCCTCTTCTCAAGTCTGTGAGGT 30534
QY 212 ----- 212
Db 30533 CCGTGTCCCTTTTGATAGATGAGGAGCTGAGACACAAAGAGTTAGTGACTTCCA 30474
QY 212 ----- 212
Db 30473 TGGCCACACAGCAGAAATGACCATAGTACAGGCCCTGTACTGGAGAAAGGTGG 30414
QY 212 ----- 212
Db 30413 GGGCGAGCCCGAGGTGGGGGAGGTGTTCAGAAACCCATCCCTCTCTGCCCCC 30354
QY 213 --GluMetAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIleSer1 232
Db 30353 AGGAGATGAACACTTGGTGTCTCAGTGTGACACCTGTATGCTAGACGTGATTAATA 30294
QY 232 LeProLeuValTyrThr----- 237
Db 30293 TCCCACTGGTGTATACACAGGTAGAGACTAGGCTGTGTGAGGCTGCCCTTTTGGAAACTG 30234
QY 237 ----- 237
Db 30233 AGGCTAGAGAGCAGCAAGAAAGCAGCTGGGTGGAGAGGCTCACCTAGAGCTAAGTGGC 30174
QY 237 ----- 237
Db 30173 TCCCTGTGAGAGTTGGGTCCACACTTTGAAGTTGGGTCTGAGACTTTGAAGTCCAAAGTTCT 30114

QY 237 ----- 237
Db 30113 AAGAGTCCAGGCTCTGCTGCGCCAGTCCAGTAGAGGCAATGTGATTATCCCATATT 30054
QY 237 ----- 237
Db 30053 AAGAGAGTTGGCGGGGACAGTGGCTCATGCTGTATTCACAGCACTTGGGAAGCTGA 29994
QY 237 ----- 237
Db 29993 GCGAGGTGATCAGCTGAGGTCAAGAGTTGAGACCAAGCTGGCCAAATGTGAAACC 29934
QY 237 ----- 237
Db 29933 CATCTCTAGAAATATACAGAAATTAGCTGTGTGTGTGACAGCCCTGTATCCAGCTAC 29874
QY 237 ----- 237
Db 29873 TTGGAGGCTGAGCGAGAGAAATCGCTTGACCCGGAGGTGAGGTGCAGTAGCTGA 29814
QY 237 ----- 237
Db 29813 GATCATGCCACTGCACCTCCAGCTGGGCGACACAGCAGACTGTCTCAACAAACAA 29754
QY 237 ----- 237
Db 29753 CAACAAACAAACAAAGGGGTTAACAGAGCCCTTAAGTCATAGTGTGCAAGTC 29694
QY 237 ----- 237
Db 29693 AGAACAGGCTTGGTCTCTGTCTCAGACTCCAGCCCTGGAGCAATCTGATTTCAGG 29634
QY 238 ----- GlnValVal 240
Db 29633 GTTCCACCTAGCCCTTGTGTACACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 29574
QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 29573 ACTGTGGGCGGTGTACAGCTCTCTCTCTGACTGTCTAGTTGGGCGGAGTTTCTGAACCA 29514
QY 261 AlaLysAlaTyrProGlyHisGluLeuAsnLeuValAlaProValPheThrPheLeuGln 280
Db 29513 GCCAAGGCTTACCTGCGCATGAGCTGAGACCTCGTTGTGCCCTTCTCAGCTTCTGCAG 29454
QY 281 PhePhePheTyrValGlyTyrPheLeu----- 288
Db 29453 TTCTTCTCTATGTTGGCTGCTGAAGGTGGGCTCTTCCAGGCGCCCTGCTGGAGG 29394
QY 288 ----- 288
Db 29393 CATGGCAGAGGGTCAATGGCCAGCAGCTGCTGAGACAGAGATGCAAGTGTCAAGAAAG 29334
QY 288 ----- 288
Db 29333 AAGGTTCACGGGTAGAAAGCAGCAGCCGTGTGGCGCACACTGTAAATCCAGCTACT 29274
QY 288 ----- 288
Db 29273 CCGGAGGCTGAGCGAGAGATGCTTGAACCCGGAGGGGAGGTTGTGTGATTGAG 29214
QY 288 ----- 288
Db 29213 ATGTGCCACTGCATCCAGCTGGGCGAAAGATGAACTCTATCTCAAAACAAACAC 29154
QY 288 ----- 288
Db 29153 AACACAAACAAAGCCCTAAGTTTCAAGAAAGCCCTTGTAGAAAGCAGAGCGGCAC 29094
QY 288 ----- 288
Db 29093 CACCTCTCTATTAACAATGCTGTGGGCTGTCTTGTCTCTCACCTCAAGTGGCTGT 29034

```
repeat_region /rpl_family="SVA" 43312..43457
repeat_region /rpl_family="Alu" complement(43963..44217)
repeat_region /rpl_family="L1" complement(45296..45533)
repeat_region /rpl_family="SVA" complement(45300..45401)
repeat_region /rpl_family="LTR5" complement(45407..45533)
repeat_region /rpl_family="LTR5" complement(45542..45818)
repeat_region /rpl_family="Alu" 45598..45770
repeat_region /rpl_family="SVA" complement(45862..46004)
repeat_region /rpl_family="L1" complement(45998..46282)
repeat_region /rpl_family="Alu" 46022..46260
repeat_region /rpl_family="SVA" complement(46692..47583)
repeat_region /rpl_family="Alu" 46885..47215
repeat_region /rpl_family="SVA" 47363..47571
repeat_region /rpl_family="SVA" complement(47665..47773)
repeat_region /rpl_family="Alu" complement(47792..48077)
repeat_region /rpl_family="Alu" complement(48110..48388)
repeat_region /rpl_family="L1" complement(48827..49102)
repeat_region /rpl_family="Alu" 48883..49110
repeat_region /rpl_family="SVA" complement(49395..49626)
repeat_region /rpl_family="Alu" 50184..50437
repeat_region /rpl_family="Alu"
```

Alignment Scores:
Pred. No.: 1.68e-132 Length: 142092
Score: 1652.50 Matches: 529
Percent Similarity: 20.318 Conservative: 0
Best Local Similarity: 20.318 Mismatches: 5
Query Match: 52.96% Indels: 2071
DB: 9 Gaps: 7

US-09-622-964-3 (1-585) x AF139813 (1-142092)

```
QY 51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluLysLeuThrLeuTyrCysAsp 70
DB 32633 AGGCTGGCCCTCAACGAGAAACAAGCTGATGTTGAGAACTGACTGTGATGAGG 32574
QY 71 SerTyrIleGlnLeuIlePheProIleSerPheValLeu 82
DB 32573 AGCTACATCAAGCTCATCCCATTTCTTCGTGCTGGTGAGTTCCCTTCGTGGCTGTT 32514
QY 82 82
DB 32513 CCGGTCCTGTGTGGCCGCCAGGCTCAAGACAGGCCAGGGAGGATCAGAGAGCTGGC 32454
QY 82 82
DB 32453 GCAAGGGGCTGGGAGGGGGCGGGAGACGCCAGCGGCAAGTGGCGCTCTCTGTAGGG 32394
QY 82 82
DB 32393 AAAGTGGGAGCTGCAAGCCAGAGAAACTGAATTAGACGTTAGGTAGACGTCCTCCGT 32334
QY 82 82
```

```
DB 32333 TAGCAATGAAAAACCCCATTTTCTAGGGAGAGGCTGACATCATGTGCTCCGTGAGCCCG 32274
QY 82 82
DB 32273 CCGGGAGAGGAGGGGGTGTGGCGGATTTCTGGGACCAAGAGGGGACCCCGGGTGACA 32214
QY 82 82
DB 32213 GAACCCCTTGGGGCTCTCGCGGCTCCATCGAGGCTCTGCTGCTCTCCGACGCC 32154
QY 82 82
DB 32153 CTTCAGAGGGGCTGGGGGCTAGGCCCGCTGCAGACAGAAAGCTGAGAGCCAGGCAT 32094
QY 82 82
DB 32093 CGCCGGGCGCTGGGCCCTGGGCTGTGGCCGAGCCCTGCCCCCTGCCCCCGCCG 32034
QY 83 83
DB 32033 CCGTCCTGCCAGGCTTCTAGCTGACGCTGCTGACCCCTGCTGGAGAACAGTACAG 31974
QY 99 AsnLeuProThrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
DB 31973 AACCTGGCGTGGCCGACCGGCTCATGAGCTGTGGGCTTCGTGGAAGCAAGGAC 31914
QY 119 GluGlnSerArgLeuAlaArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
DB 31913 GAGCAAGCCGCGCTGCTGCGGCGGACCTCCGCTACGCAACCTGGGCAACGCTGCTC 31854
QY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnIleVal 158
DB 31853 ATCTGGCGAGCGTCAGCACCGGCTGACAAAGGCTTCCGACGCGGACGACTGGTG 31794
QY 159 Gln 159
DB 31793 CAAGCAGGTGGCGGAGCGGAGGACMACGGGAGGACCGGCGAGGCCGAGAT 31734
QY 159 159
DB 31733 GGGCGCGGCAAGTAAGTAAGTGGTGGAGCCAAAGTCCCGGACTCGGGGATTGGGT 31674
QY 159 159
DB 31673 GAGGCCAGAGTGGGGTGTGTCAAAGATTGGGGTCCAAATTGGCGGAGACAGTGGG 31614
QY 159 159
DB 31613 TGCTGAAGTGGGGCGAGGCCAGAGGCCACCCCTCGAGAGTAGAGTGTGAGGACGG 31554
QY 159 159
DB 31553 CTAAGGACCTTGAAGGATTAATGAAGAAGAGGAGGCTTGGGAACGTGAGTACT 31494
QY 159 159
DB 31493 AGGGTACTTCCCTGTGCCCTTGCCCTTGATGTCGCGTTTCCACTCTGAGAGTATG 31434
QY 159 159
DB 31433 GGAATTTGGTCTGTGACACCCCTCAAGCTGAGCTGACCTGCTGTTAATTAAGACAG 31374
QY 159 159
DB 31373 ACCCAGGCTAGGGCTGTGGCTCTGCGCTGATATCCAGTGTCTTAGGAGCAAGGTGG 31314
QY 159 159
DB 31313 GAAGATCGTGTGAGCCAGCTGTTTGAAGGCCCTGAGACACATAGCGAGACCCCATC 31254
QY 159 159
```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Bergen, A.A., McGarty-Dugan, V., Figueroa, D., Austin, C.P.,
Metzger, M.L., Caskey, C.T. and Madellis, C.
Identification of the gene responsible for Best macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
98324772
9662395
2 (bases 1 to 142092)
McGarty-Dugan, V.A., Hammond, H.A., Clement, M.K., Larson, D.R.,
Liu, X., Soderman, A.R., McGowan, J.M., DeAngelis, D.M., Lin, C.,
Fitzpatrick, E.S., Harrison, K.M., Petraklin, K., Caskey, C.T. and
Metzger, M.L.
Direct Submission

TITLE
JOURNAL
REMARK
Submitted (01-APR-1999) Department of Human Genetics, Merck & Co.,
Inc., Summerville Pike, West Point, PA 19486, USA
IMPORTANT: This submission contains the entire insert of clone
pJ759j12 which comes from a PAC library constructed at the Roswell
Park Cancer Institute by the Pieter de Jong group. This clone has
been finished according to strict quality criteria, and attempts
have been made to resolve all base calling problems such as
compressions and repetitive elements. This sequence has been
finished such that all consensus base calls consist of two or more
separate clones with double-stranded coverage or two or more
separate clones with two types of sequencing chemistry. The
expected Phred/Phrap calculated errors/10Kb is 0.00.

FEATURES
Location/Qualifiers
1. 142092

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="pJ759j12"
complement(19..1855)
repeat_region
/rpt_family="L1"
1861..2001
repeat_region
/rpt_family="L1"
1990..2317
repeat_region
complement(2031..2257)
/rpt_family="SVA"
2320..2974
repeat_region
/rpt_family="L1"
2985..3540
repeat_region
/rpt_family="Alu"
complement(4034..4124)
repeat_region
/rpt_family="MIR"
4417..4995
repeat_region
complement(4432..4936)
/rpt_family="SVA"
complement(5503..6111)
repeat_region
/rpt_family="Alu"
5569..6095
repeat_region
/rpt_family="SVA"
complement(7239..7542)
repeat_region
/rpt_family="Alu"
7306..7497
repeat_region
/rpt_family="SVA"
7752..7828
repeat_region
/rpt_family="MIR"
9099..9403
repeat_region
/rpt_family="Alu"
complement(9526..9897)
repeat_region
/rpt_family="MER1"
9635..9752
repeat_region
/rpt_family="MER1B"
11924..12197
repeat_region
complement(11955..12151)
/rpt_family="SVA"
complement(12767..13466)
repeat_region
/rpt_family="Alu"
16757..17039
repeat_region
/rpt_family="Alu"

repeat_region
complement(16775..17010)
/rpt_family="SVA"
repeat_region
complement(18220..18507)
/rpt_family="Alu"
18229..18485
repeat_region
/rpt_family="SVA"
20101..23446
/gene="HUMFERRITIN"
/note="Human ferritin heavy chain gene"
join(20101..20421,22218..22364,22621..22746,22842..23446)
/gene="HUMFERRITIN"
/product="ferritin heavy chain"
complement(21466..21731)
repeat_region
/rpt_family="Alu"
21506..21714
/rpt_family="SVA"
complement(23272..37375)
/gene="YMD2"
/note="Best macular dystrophy gene"
complement(join(23272..23629,24840..25478,27690..27841,
28155..28235,29427..29579,30274..30351,30740..30894,
31787..32020,32537..32631,35777..35964,37308..37375))
/gene="YMD2"
/product="Best macular dystrophy protein"
complement(23735..23954)
repeat_region
/rpt_family="Alu"
23789..23956
repeat_region
/rpt_family="SVA"
26429..26707
repeat_region
/rpt_family="Alu"
complement(26444..26696)
/rpt_family="SVA"
27103..27349
repeat_region
/rpt_family="Alu"
complement(27106..27323)
/rpt_family="SVA"
repeat_region
complement(29161..29307)
/rpt_family="Alu"
repeat_region
complement(29762..30043)
/rpt_family="Alu"
29786..30017
repeat_region
/rpt_family="SVA"
30462..30516
repeat_region
/rpt_family="MIR"
complement(31095..31409)
repeat_region
/rpt_family="Alu"
33005..33260
repeat_region
/rpt_family="Alu"
33247..33597
repeat_region
/rpt_family="L1MC2"
complement(33849..34431)
repeat_region
/rpt_family="Alu"
34076..34411
repeat_region
/rpt_family="SVA"
complement(34640..34890)
/rpt_family="Alu"
34972..35401
repeat_region
/rpt_family="Alu"
complement(34988..35392)
repeat_region
/rpt_family="SVA"
36238..36487
repeat_region
/rpt_family="Alu"
complement(36253..36457)
repeat_region
/rpt_family="SVA"
complement(38661..38855)
/rpt_family="L1"
repeat_region
complement(40997..41224)
/rpt_family="Alu"
41056..41213
repeat_region
/rpt_family="SVA"
42745..43020
repeat_region
/rpt_family="Alu"
complement(42763..42966)

Db 90192 CAAGGGTTCTTCCAGGGATCCAGTCTGTCTCGCTTCTTTCTTTTCTTTT 90251
 QY 366 ----- 366
 Db 90252 TAAAGGAGTTCTACTCTTTGTTGCCAGAGCTGGAGTGCAGTAACTCGGCTCAC 90311
 QY 366 ----- 366
 Db 90312 TGCACCTCCGGCTCCAGATTCAAGCAATTTCTCTGCTCAGCTCTGAGTACGTGG 90371
 QY 366 ----- 366
 Db 90372 ATTATAGTGGCCAGCCACCAGCCGGCTAATTTTGTATTTTACTAGACAGTTTCA 90431
 QY 366 ----- 366
 Db 90432 CCATGTTGGCCAGGCTGGTCTGAACTCCTGACCTGAGTGATCCACCCTCTCAGCCTC 90491
 QY 366 ----- 366
 Db 90492 CCAAGTGTGGGATTACATGTGTGAGCCACTGTGCTGGCTGCTTTCTTTTAAAGC 90551
 QY 366 ----- 366
 Db 90552 CAAATATCTACTAGACTGCAATCGAGTTTAACTACAGTCTATAGTAGTGTGAGGAAG 90611
 QY 366 ----- 366
 Db 90612 GTTGGGAAGTGATCAAAATGAAGCTGAGGCTTGTAGTCAAAATTTCTGAGG 90671
 QY 366 ----- 366
 Db 90672 ATGACTTTGAGCCCTACATGGTCTGTACCCAGCAGTGAAGTTGTGAGGGGTGGGA 90731
 QY 366 ----- 366
 Db 90732 GGGCTGAAGAACAGATGAAGCATAGACCTTGTCTCAAGAGATGCACATTTATGGA 90791
 QY 366 ----- 366
 Db 90792 GGGAGCTCAACCCCAAGTCTCAAACTCTGATACAGAGTACAAAGTACTGATGTCCAGA 90851
 QY 366 ----- 366
 Db 90852 AAAGGACAGAACTGAAACACAGTCACTTTGTCTGCTGGAGCGGCTTCCAGCTGG 90911
 QY 366 ----- 366
 Db 90912 GTCTGAGCTGAGCCATGGAACATGGAGAAATCTGAATTTGGGCAAGGCGAGCCATAC 90971
 QY 366 ----- 366
 Db 90972 TCTCTGTAGTAAAGCTTTCCCTTGCAGAGGTAAAGTCTGGGGCTCCCGGATCCCTGTT 91031
 QY 366 ----- 366
 Db 91032 GCTAGGAAGTCAAAATTTCTTTGTGTGATGTCACTCCAGTTGGAACACAATTTCTG 91091
 QY 366 ----- 366
 Db 91092 CATTGCCAGAGTCACTCATGGGCTCATGTGAACAATCTATGCCAGGCGACAGTGT 91151
 QY 366 ----- 366
 Db 91152 CTGACTCCCTGAGTGAAGGGTTTACAGGGGAAGTAATGATGAGAGGCTTTACACG 91211
 QY 366 ----- 366
 Db 91212 CCAGCGGGGGTGGTGGGGGGTTGATGTTAACTGTGTCAAGAGGAATCAACAACA 91271
 QY 366 ----- 366

Db 91272 GTGAGGTGAGCTGGGCTGGAGGATCACCGGAGGTACAGGACAGATCAGAGAGGT 91331
 QY 366 ----- 366
 Db 91332 GAGAGCTGGGGCATGTGTAGGAAGACGTGTGGCTTGGCTTGGCCAACTAGAGAGAG 91391
 QY 366 ----- 366
 Db 91392 GAGCGGGGTAAAGGAGAAATAGGCCAGGTGTGCTCTTGTCCAGTGGCTCAGCCCT 91451
 QY 367 ----- SerLeuAsnLysGluGluMetGluPheGluProAsnGlnG 380
 Db 91452 GCATGCTCCTGTTCTTTCCAGCTGACACAAAGAGATGGAGTTCCAGCCCAATCAGCA 91511
 QY 380 uAspGluLysAlaHisAlaGlyTLeileGlyArgPheLeuGlyLeuGlnSerHisAs 400
 Db 91512 GGACGAGGAGATGTCTACAGCTGGAGTATGTGGCGCTCTCCTAGGCTCAGTCCCATGA 91571
 QY 400 PHisHsProProAlaAsnSerArgThrLysLeuLeuProLysArgGluSerLe 420
 Db 91572 TCACCATCTCCAGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAAATCCCT 91631
 QY 420 uLeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnG 440
 Db 91632 TCTCCACGAGGCGCTGCCCAAAACACAGGACAGCCAAACAGAACTTTAGGGGCCAGGA 91691
 QY 440 uAspAsnLysAlaThrLysLysLysAlaValAlaAspAlaPheLysSerGlyProLeuTyrG 460
 Db 91692 AGACAAACAAGGCTGGAAGCTTAAAGCTGTGGAGCCCTTCAAGTCTGCCCCACTGATACA 91751
 QY 460 nArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrProMetPhePhePr 480
 Db 91752 GAGGCCAGCTACTACAGTGTGCCACAGACACCCCTCAGCCCTCCATGTTCTTCC 91811
 QY 480 oLeuGluProSerAlaProSerLysLeuHisSerValThrGlyTLeaPThrLysAspLys 500
 Db 91812 CCTAGAACCATCAGGCGCGCTCAAGCTTCACAGTCTCAGGCAATGACCAAGACAA 91871
 QY 500 sSerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAs 520
 Db 91872 AAGCTTAAAGACTGTGAGTCTTGGGGCCCAAAAGTTTGAATGTCTCAGAGACGGA 91931
 QY 520 pGlyAlaLeuMetGluHisProGluValSerGluValArgArgLysThrValGluPheAs 540
 Db 91932 TGGGGCTTGTGAGACACCCAGAAATATCTCAAGTGTGGAGAAACTGTGGAGTTTAA 91991
 QY 540 nLeuThrAspMetProGluLeuLeuProGluAsnHisLysGluProLeuGluGlnSerPr 560
 Db 91992 CCTGACGGATATGCCAGAGATCCCGAAATACACTCAAAAGACCTTTGGAACATATCAC 92051
 QY 560 oThrAsnLysHisThrThrLeuLysAspHisMetAspProTyrTrrPalaLeuGluAsnAr 580
 Db 92052 AACCAATATACACTACACTCACTCAAAAGATGATGATCTTATTTGGGCTTGGAAACAG 92111
 QY 580 gAspGluAlaHis 584
 Db 92112 GTCTGTCTCTCAC 92124

RESULT 6
 AF139813/c 142092 bp DNA linear PRI 22-APR-1999
 LOCUS Homo sapiens clone pdu759j12 chromosome 11 map 11q13, complete
 DEFINITION
 ACCESSION AF139813
 VERSION AF139813.1 GI:4633837
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Petukhin,K., Kolstel,M.J., Bakall,B., Li,W., Xie,G., Marknell,T.,
 Sandgren,O., Forsman,K., Holmgren,G., Andreasson,S., Vujic,M.,

QY 288 ----- 288
Db 88032 CCTTCCTACCTCTCTTATTTTGGTAATGGGGGTAGTCTCTCTCTGCCCTTCC 88091
QY 288 ----- 288
Db 88092 TGTCACTGTGACACACACACACACACACACACACACACACACATTCCTAT 88151
QY 288 ----- 288
Db 88152 TCCTCTAAATTTCCCTCCCTCCCTCCCTATCTTGGTCTGACATCAAAACAAATCA 88211
QY 288 ----- 288
Db 88212 CACTTTATGCTTGAATTTCCAGGGTCCCTCCAGTGGCTGCAAGATGTCCTCGACC 88271
QY 288 ----- 288
Db 88272 CCTAAGCAGACCGGTGTACCTCTCTCGGGCTTTGTAGGGCATTTTAGAGTTGCTAT 88331
QY 288 ----- 288
Db 88332 CCAGGAATCTGCCACCTTAGACTGCCCTTAGTTAGCCAGCTTCAATATATCTCTG 88391
QY 288 ----- 288
Db 88392 TTGCATGAATGAATAAATTATGCAACTCCAGGTAGATACATGAGGTAGATAAGGCA 88451
QY 288 ----- 288
Db 88452 GTGACTACGCGAGTGATACACTCAGAGGACAGCTGTGGTGTTCAGGAAGAGCTGGCTC 88511
QY 288 ----- 288
Db 88512 AGAAGATTAGAGGGGCTGTGTCCAGAAGTGTGGGTCCCAAGATGTGGGGGCTGG 88571
QY 288 ----- 288
Db 88572 AGCCCTAAACTGCTTGTGAAGACAGTGTGTCAGGAGGAAGGCTTCATGGGGTGTGA 88631
QY 288 ----- 288
Db 88632 AATACGACAGCTGAGGTTTAAAGGGGGAAGCTGCTTGAAGAGTCTGCTGAGGCT 88691
QY 289 ----- 301
Db 88692 TACAGAGCCTCAGCTGTCCCAAGGTGGCAGAGCACTCATCAACCCCTTGGAGAGAT 88751
QY 302 AspaPAPpPhegluThrAsnTrpIleValAspArgAsnLeu ----- 315
Db 88752 GATGATGATTTTGAACCACTGATGTGTGACAGGAATTTGACAGTATGGGGAGAGGA 88811
QY 315 ----- 315
Db 88812 GAGAAACCATACATGACCTTCCCAAAAGTGACCCAAAGAGAGAGACCACCTGTCTG 88871
QY 315 ----- 315
Db 88872 TAGGAGGCTCAGACGATGATCAACCTTCCCTCTCTCTCTCAGCAGTCAAT 88931
QY 315 ----- 315
Db 88932 CACTCAGAGATTTCTACCTCAATCTTTGAGGCTGACAGGACAGCACCATCTCCCATTT 88991
QY 315 ----- 315
Db 88992 CACAGCAGGGAACCTAGGTCAGAGAGAGAGAGAGATTCCTCAAGTATCAGGACA 89051
QY 316 ----- 317
Db 89052 TACAAAGTCTGCTGGAGATGATCTTCTGTGGAGCTTCTTCTCTCTCTGTCACAGGT 89111

QY 317 lserIleuAlaValAspIleuMetHisGlnAspLeuProArgMetGluProAspMetTy 337
Db 89112 GTCCCTGTGGCTGTGATGATGACACAGGACCTCTCGATGGAGCGGACATGTA 89171
QY 337 rTTPAsnLysProGluProGlnProProtyrThrAlaIleSerAlaGlnPheArgArgAl 357
Db 89172 CTGGAAATTAAGCCGAGCAGACAGCCCTTACACAGTGTCTTCGCCAGTTCGCTGAGC 89231
QY 357 aserPheMetGlySerThrPheAsnIle ----- 366
Db 89232 CTCTTTATGGGCTCCACCTTCAACATCAGGTGTGGCCAGAGCGGGGCTGTGGGA 89291
QY 366 ----- 366
Db 89292 AGCCCTCTAGTGACAGGGGTGCTGCTAGGAACTTAGAATAGCACTAGTAATGATACA 89351
QY 366 ----- 366
Db 89352 GGTGCTTCAAGTAAGTGACAGGCACTACTATGCTTTATTAACATTAATATTTT 89411
QY 366 ----- 366
Db 89412 TCCTCCAAATTAATCTGTTGTGTATCCCAAGTTTCAGATTAATTAAGTACAGTTCA 89471
QY 366 ----- 366
Db 89472 AGAGAGTAAGTTGCCAAGGCCACATAGCTACCAAAATGTGCAATTTGCTACTGAAAGAC 89531
QY 366 ----- 366
Db 89532 AGCCTATGATCAGTGATGATGATGGAACGTTAGAGCTGCTGTGTATCAGAACTATG 89591
QY 366 ----- 366
Db 89592 TTTTCTTTCTTTTGTAGACATATCTGCTGTGTGCGCCAGTTGGAGCGAGTGGCT 89651
QY 366 ----- 366
Db 89652 GATCTTGCTACTGCAACCTCCGCTCTGTGGTTCAAGTATCTCTGTTCAGGCTC 89711
QY 366 ----- 366
Db 89712 CCCACTACCTGGATTAAGGTGCCACACACACACTGCTAATTTTGTACTTTTAACT 89771
QY 366 ----- 366
Db 89772 AGAGATGAGTTTACCATGTTGGCAGGCTGTCTCCAACTCTGACACAGTAATCTGCC 89831
QY 366 ----- 366
Db 89832 CGCTTGGCTCCCAAAATGCTGAATTAAGTGTCAAAACTATGTTTCTGATTAAGCT 89891
QY 366 ----- 366
Db 89892 ACATGCTTGGATGGAAAGTGGAGTGGGTTCCCTGGGATGGGGAGGGCAGCAAACT 89951
QY 366 ----- 366
Db 89952 CCCAGACGACGACGACCATCAGAGTACTCTGAATGACTTGTGCTAAGAGATA 90011
QY 366 ----- 366
Db 90012 AGGGCTCAGGCAACCCACAGCAGCAGCACTTATCCCAATGTCCTCACTTCCCTGATTC 90071
QY 366 ----- 366
Db 90072 CATCTGAATCCCTCTTAGCTGAGTGGGCTGAAGGGCTATCCAGCTGTCTTTCTCC 90131
QY 366 ----- 366
Db 90132 CCAGACACAGAGTTGAAGTGCCTTGGAGAGTGTGGGCACATGTACAGGTTCACTACT 90191
QY 366 ----- 366

Db	85812	TAAGATCGACCCGCTGACTCCAACCTCGGTGACAGAGCCAGACCCCTTTCTGGAATA	85871
QY	159	-----	159
Db	85872	AATTAATACCTGCCCCACATGCTCAGGCCAGAACAGCAGCTAGTAGTCTCAGAAATT	85931
QY	159	-----	159
Db	85932	TTTTGTGTGAAGAAAGAGATGGCAAGAGAGTGTGAGTTCTTATAGTCAGCAGG	85991
QY	160	-----Ala1 161	161
Db	85992	TGCGGCCATCCCTTCTGAGGTTCTCCACCCAGCCCTTCTTCACTCCACTCTGACG	86051
QY	161	-----	161
Db	161	YPHEMETHPRoAlaGluHisGlyGlnLeuGluIuylsLeuSerLeuProHisMetPh	181
Db	86052	CTTTATGACTCCGGCAACACACAGCAAGTTGGAAACTGAGCCCTACACACAACTGTT	86111
QY	181	eTTPValProTPValTPPheAlaAsnLeuSerMetLysAlaTPLeuGlyValArg11	201
Db	86112	CTGGTGCCCTGGGTGTGGTTGGCCAACTCTCAATGAAGCGTGGCTTGGAGTGAAT	86171
QY	201	eArgAspProIleLeuLeuGlnSerLeuAsn-----	212
Db	86172	CCGGGACCTATCTGCTCCAGAGCCTGCTGAACGTGAGCCACCTGTACAGACAGGCTG	86231
QY	212	-----	212
Db	86232	CCGCAAGTGGAGAGGTTGTGTCCACAGAAACAGGTTCTTCAAAAGAGAGCTT	86291
QY	212	-----	212
Db	86292	GGGCCCTGAGGGTCTTCCGAGAGCGGAGGTGGGTTGCAGAACTTTTCCACAGCAA	86351
QY	212	-----	212
Db	86352	TCCACAGCCCGAGGTGTCCTTCTCAGAGCCCTCTCTTCCAAAGTGTGAGGT	86411
QY	212	-----	212
Db	86412	CCTGTTCCCTTTTATAGATGAGAGAGCTGACACAAAGAGTTTACTGAGCTTCCCA	86471
QY	212	-----	212
Db	86472	TGGCCACACAGCAGAGATGACATAGGTACAGGCCCTGTACTGAGAGAAGTGG	86531
QY	212	-----	212
Db	86532	GGGCGAGCCAGAGGTGGGGGAGGTGTTCAGAACCCATCCCTCTTCTGCCCCC	86591
QY	213	--GluMetAsnThrLeuArgThrGlnCysGlyHisLeuTyralaTyraAspTrpIleSer1	232
Db	86592	AGGAGATTAACACCTGTGCTACTGAGTGTGACACCTGTATGCTTACGACTGATTA	86651
QY	232	LeProLeuValTyThr-----	237
Db	86652	TCCCACTGGTGTATACACAGGTGAGACTAGAGCTGTGAGGTGCTTTTGGGAAACTG	86711
QY	237	-----	237
Db	86712	AGGCTAGAGACCAAGAGAGAGCTGGGTGGAGAGGCTACCTAGAGCTAAGTGGC	86771
QY	237	-----	237
Db	86772	TCCCTGGGAGTTGGGTCCACACTTTGAAGTTGGGTCTGACTTTGAAGTCCCAAGTTCT	86831
QY	237	-----	237
Db	86832	AAGAGTCAGGCTCTGCTGCTGGCCAGTCCAGTAGAGGCAATGTATATCCCATATTA	86891
QY	237	-----	237

Db	86892	AAGAGAGTTGGCCGGGACAGATGGGTCTCATGCTTATATCCAGCACTTTGGGAAGCTGA	86951
QY	237	-----	237
Db	86952	GGCAGGTGATATCACTGAGGTTCAGAGTTTGAGAACCAAGCCTGGCCAAATGTGAACCC	87011
QY	237	-----	237
Db	87012	CATCTCTAGTAATAATACAGAAATTAGCTGTGTGTGTGCACGCTGTAAATCCAGCTAC	87071
QY	237	-----	237
Db	87072	TTGGAGGCTGAGGCGAGAGAAATCGTTGAACCCGGGAGGTGAGGTTGAGTAGCTGA	87131
QY	237	-----	237
Db	87132	GATCATGCCACTGCTCAAGCTTGCGGCGACACAGCAAGACTGTCTCAACAAACAA	87191
QY	237	-----	237
Db	87192	CAACAAACAAACAAACAAACAAACAAAGGGTTAACAGAGCCCTAAGTCACATTAAGT	87251
QY	237	-----	237
Db	87252	TGCAAGTCAGAACAGAGCCTTGCTCTCTGTCTCAGACTCCAGCCCTGAGACATCTG	87311
QY	238	-----G 238	238
Db	87312	ATTTCAGGGTCCACACTAGCCCTTGTGTACCAACATCTCTCTCTCTCTCTCTCTCC	87371
QY	238	InvAlValThrValAlaValIlyrSerPhePheLeuThrCysLeuValIglyArgGlnPhe1	258
Db	87372	AGGTGTGACTGTGGCGGTGTACAGCTTCTCTGCTGTAGTTGGCGGCAAGTTTC	87431
QY	258	euaSProAlaLysAlaTyProGlyHisGluLeuAspLeuValIProValPheThrP	278
Db	87432	TGAACCCAGCCAAAGGCTTACCTGGCCATGAGCTGACCTGCTGTGGTCCGCTTACAGT	87491
QY	278	heLeuGlnPhePhePheTyValIglyrPleu-----	288
Db	87492	TCTCAGATTCTTCTTATGTGTGTGCTGCTGAAGGTGGGCTCTCCAGAGGCCCTGCTG	87551
QY	288	-----	288
Db	87552	GCTGAGCGATGGCCAGAGGGGTATGGCCACAGCTGCTGAGAGAGATGACAGTGC	87611
QY	288	-----	288
Db	87612	AGGAAGAGAGTCTCACGGGTAGAAAGCAGCAGCCGTGTGGCCACACTGTATATCC	87671
QY	288	-----	288
Db	87672	CAGTACTCGGAGAGCTGAGCAGAGATGCTTGAACCCGGAGGGGAGTGTGTGT	87731
QY	288	-----	288
Db	87732	GAGTTGAGATGTGCCACTGCATCTCCAGCTGGGCAAAAAGATGAATCTATCTCAAA	87791
QY	288	-----	288
Db	87792	ACAACAAACAAACAAACAAAGCCCTAAGTTTCAGAAAGCCCTGCTTAAAGAGCG	87851
QY	288	-----	288
Db	87852	AGCGGACACACTCTCTTATTAAGATGCTGTGGGCTGTCTGTCTACCTCAAG	87911
QY	288	-----	288
Db	87912	TGGCTGTCCAGGATTCCTCCCAACACAGCAATACTCCGAACAGATGTTCTGAATCA	87971
QY	288	-----	288
Db	87972	CACAGTTTCTCCACCTCTTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC	88031


```

repeat_region 71536.71637
/rpl_family="LFR5"
repeat_region 72720.72948
/rpl_family="L1"
complement(73488.73633)
/rpl_family="Alu"
complement(73925.74200)
/rpl_family="Alu"
repeat_region 75721.75955
/rpl_family="Alu"
repeat_region 77185.77449
/rpl_family="Alu"
repeat_region 78093.78222
/rpl_family="L1"
repeat_region 78227.78294
/rpl_family="L1"
complement(80458.80707)
/rpl_family="Alu"
complement(81544.81974)
/rpl_family="Alu"
repeat_region 82055.82305
/rpl_family="Alu"
repeat_region 82514.83096
/rpl_family="Alu"
complement(83685.83940)
/rpl_family="Alu"
repeat_region 85336.85850
/rpl_family="Alu"
complement(86429.86483)
/rpl_family="MIR"
repeat_region 86802.87183
/rpl_family="Alu"
repeat_region 87646.87792
/rpl_family="Alu"

```

Alignment Scores:

```

Pred. No.: 6 74e-133 Length: 112309
Score: 1655.50 Matches: 529
Percent Similarity: 20.29% Conservative: 0
Best Local Similarity: 20.29% Mismatches: 5
Query Match: 53.06% Indels: 2073
DB: 9 Gaps: 7

```

US-09-622-964-3 (1-585) x AC003025 (1-112309)

```

QY 51 ArgLeuAlaLeuThrGluGluGlnLeuMetPheGluIleuThrLeuThrLeuTyrcysasp 70
|||||
DB 84312 AGCGTGGCCCTCAGCGAAGACACACGCTGATGTTGAGAAACTGACTCTGTAATGCGAC 84371
71 SerTyrIleGlnLeuIleProIleSerPheValLeu-----
|||||
DB 84372 ACCTACATCCAGCTATCCCATTTCTTCCTGCTGGTGAGTTCCTCCCTCTCGGCTGT 84431
82 -----
QY 82 -----
DB 84432 CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGCCAGGGAGGATCAGAGAGAGCTGCG 84491
82 -----
DB 84492 GCAGGGGCTGGGAGGGGGCGGGGAAAGCCAGCGGAGTCCGCGCTCTCTAGAG 84551
82 -----
QY 82 -----
DB 84552 AAAGTGGGACTGCAGCCAGAAACTGAAGTTAGAGTTAGAGACGTCTGCGGT 84611
82 -----
QY 82 -----
DB 84612 TAGCAATGAAACCCATTCTCTGAGGGAAGCGCTGACATCATGTCTCTGAGCCCTG 84671
82 -----
QY 82 -----
DB 84672 CGCGGAGGGAGGGGGTCTGGCGATTCTGGGACAGCAGAGGGAGACCCCGGGGTGACA 84731

```

```

QY 82 -----
DB 84732 GAACCTTGGGGCTCTGCGCCCTCCATGCGAGGCTCTGCTCTGCTCCGAGGCC 84791
82 -----
QY 82 -----
DB 84792 CTTCAGAGAGGCTGGGGGCTAGCCCGCTCCAGAGAAAGCTGAGAGCCGAGCAT 84851
82 -----
QY 82 -----
DB 84852 CGCCGGCGCTGGCCCTGGGCTGTGGCCGAGCTGCCCCCTCGCCCCCGCC 84911
83 -----
QY 83 -----
DB 84912 CCTCTCTCCAGGCTCTACGTGACGCTGTGTGAGACCCGCTGTGTGAACGATACGAG 84971
99 AsnLeuProIleProAspArgLeuMetSerLeuValSerIlePheValGluGlyLysAsp 118
|||||
DB 84972 AACCTGCGGTGGCCGACCGCTCATGAGCTGTGTGCGGGCTTGTGTGAAAGCAAGGAC 85031
119 GluGlnSerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
|||||
DB 85032 GAGCAAGCGCGGCTGTGGCGGCGACGCTATCCGCTACGCCAAGCTGGCAAGCTGCTC 85091
139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnIleVal 158
|||||
DB 85092 ATCTGCGGAGCTGTGACCGACCGAGTCTACAAAGCGCTTCCCGAGCCCGACCTGTGTG 85151
159 Gln-----
QY 159 -----
DB 85152 CAAGCAGTGGGGGAGCGGAGCAACGGGAGGACACCGGACAGCCAGGGCCGAGAT 85211
159 -----
QY 159 -----
DB 85212 GGGCGCGCAGAAATGAGATGGTGGAGCCAAAGTCCCGGACTCGGGGATTGGGT 85271
159 -----
QY 159 -----
DB 85272 GAGGCCAGAGTGGGGTGTGTCAAGATTGGGGGTCCAAATTGGCGGAGAGAGTGGGG 85331
159 -----
QY 159 -----
DB 85332 TGTCTGAAGTGGGGGAGGCCAGGACCCACCTCCGAGACTAGAGTCTGAGCAGGG 85391
159 -----
QY 159 -----
DB 85392 CTAGAGACCTTGAAGGATATATGAAGAAGGGTGACGGCTTGGAACTGTGAGAGTACT 85451
159 -----
QY 159 -----
DB 85452 AGGGTCTACTTCCCTCTGCCCCCTTGCCCTCTTGATCTCCGGTTTCCACTGTGAGTATG 85511
159 -----
QY 159 -----
DB 85512 GGACATGTGTCTGACACCCCTCAGCCTGAGCTGTGCTGTGTTAATAAGACAG 85571
159 -----
QY 159 -----
DB 85572 ACCAGGCTAGCGGTGTGCTCTGCTGTAAATCCAGTCTTTAGAGGCAAAAGTGG 85631
159 -----
QY 159 -----
DB 85632 GAAATGCTTGAGCCCAAGCTGTGTTGAGACGCCCTTAGCAACAATAGCAGACCCCAATC 85691
159 -----
QY 159 -----
DB 85692 TCTACAAAACATTAATAATTAGCAGGAGCATGTGGCGTGTGCTGTAGTGTGAGGCTGA 85751
159 -----
QY 159 -----
DB 85752 GTATCGGAGGCTGAGCGAGAGGATCACTTGAAGCCAGAGATTCCAGGCTGAGTGGCG 85811
159 -----
QY 159 -----

```

TITLE
JOURNAL
Davis, C., English, C., Franklin, T.L., Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J., Hinson, S., Narayanaswamy, U., Newton, J., O'Brien, K., Oliver, T., Patel, P., Probst, S., Rayner, S., Schageman, J., Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T., and Wilson, R.
Direct Submission
Submitted (21-Oct-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
3 (bases 1 to 112309)

REFERENCE
AUTHORS
Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M., Buetner, D., Bumeister, R., Card, P., desalibcat, F., Dunn, J., English, C., Ehrhardt, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
Direct Submission
Submitted (23-Jul-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
On Jul 23, 1998 this sequence version replaced gi:2554967.

COMMENT
JOURNAL
IMPORTANT: This submission contains the entire insert of clone PDJ466611. PDJ466611 comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and as compressions have been made to resolve all base calling problems such as calculated errors/10kb is 0.18. In addition, this sequence has been finished such that 99.9% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.
Further information regarding the map of this region or annotation of PDJ466611 can be found at <http://gestec.swmed.edu/chromsol.htm>.
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST AHNAK. This region spans over 1.5 Mbp.
MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA)
MAPPED CLONE OVERLAP: HTGS submitted PAC clones PDJ319013 and PDJ756b9.

FEATURES
source
1. .112309
/organism="Homo sapiens"
/db_xref="taxon:9606"
complement(4341..4656)
/rpt_family="Alu"
repeat_region
complement(9880..10120)
/rpt_family="Alu"
repeat_region
10163..10282
/rpt_family="Alu"
complement(11232..12157)
/rpt_family="Alu"
repeat_region
13351..13560
/rpt_family="L1"
complement(15092..15146)
/rpt_family="MIR"
15826..15932
/rpt_family="MIR"
repeat_region
17011..17146
/rpt_family="MIR"
complement(18018..18313)
/rpt_family="Alu"
repeat_region
18493..18741
/rpt_family="Alu"
repeat_region
18820..18954
/rpt_family="Alu"
repeat_region
/rpt_family="Alu"
complement(21374..21661)
/rpt_family="Alu"
repeat_region
22809..23212
/rpt_family="Alu"
repeat_region
23499..23752
/rpt_family="Alu"
repeat_region
24118..24472
/rpt_family="Alu"

repeat_region
25005..25113
/rpt_family="MIR"
repeat_region
31160..31442
/rpt_family="Alu"
complement(32651..32947)
/rpt_family="Alu"
complement(35970..36025)
/rpt_family="MIR"
38958..39145
/rpt_family="Alu"
44878..45480
/rpt_family="Alu"
49769..50071
/rpt_family="Alu"
complement(50172..50310)
/rpt_family="Alu"
50902..50982
/rpt_family="MIR"
51038..51659
/rpt_family="Alu"
54494..54769
/rpt_family="Alu"
55529..55623
/rpt_family="THE1"
complement(55691..55967)
/rpt_family="Alu"
complement(56049..56257)
/rpt_family="Alu"
56721..57100
/rpt_family="THE1"
58402..58680
/rpt_family="Alu"
complement(58748..59345)
/rpt_family="Alu"
59678..59904
/rpt_family="Alu"
61149..61424
/rpt_family="Alu"
complement(61930..62242)
/rpt_family="Alu"
62276..62710
/rpt_family="MIR1"
62915..63118
/rpt_family="Alu"
complement(63349..63455)
/rpt_family="Alu"
64339..64614
/rpt_family="Alu"
complement(64707..64984)
/rpt_family="Alu"
complement(65596..65811)
/rpt_family="Alu"
complement(66500..66753)
/rpt_family="Alu"
67311..67542
/rpt_family="Alu"
67835..68119
/rpt_family="Alu"
68860..69150
/rpt_family="Alu"
69164..69272
/rpt_family="Alu"
69354..70245
/rpt_family="Alu"
70655..70939
/rpt_family="Alu"
70933..71075
/rpt_family="L1"
71119..71395
/rpt_family="Alu"
71404..71530
/rpt_family="LTR5"

REFERENCE 2 (bases 1 to 1289)
 AUTHORS Marmorstein, L.Y., McLaughlin, P.J., Stanton, B., Yan, L., Crabd, J.W.
 TITLE Direct Submission
 JOURNAL Submitted (30-NOV-2001) Cole Eye Institute, 131, Cleveland Clinic
 Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA
 FEATURES
 source Location/Qualifiers
 1..1289
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 <1..1289
 /product="bestrophin"
 /protein_id="AA140882.1"
 /db_xref="GI:18476496"
 /translation="PQHLVAKGFTPESEHLEKSLPNSFWPMVYFANLSTKAVI
 GGRIRDPVLQSLIDEMNTLTQCGHLYADWISVPLVYQVAVAYSEFLACIVR
 QFLNPAKAYGHENDLVPLFTLQFFVAGWIKVABOLINPREDDDDEETWYDR
 SLQVLSADVEMHNDLPEMERDMYNDPEHPPTAASRRSPRGSTFNSLAGE
 DMFOPEEEREAHTGILGHFLGSSDHPDPTNSKTLMPKKEGFHGHKRNKG
 ARLDSDQEDSKAMREGGFKSALCGPGCHSAPQTPLEHTPVEEAPLGLRV
 SGIDEAKDQSLQAPPSIKKSFELPESLEASAEPLQSGSHVRKTYEFMLADLSEAP
 EHLKEPNEPHEHIALKNDHPYALLENREDAHS"

BASE COUNT 293 a 390 c 360 g 246 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,1e-139 Length: 1289
 Score: 1701.00 Matches: 324
 Percent Similarity: 81.528 Conservative: 29
 Best Local Similarity: 74.838 Mismatches: 72
 Query Match: 54.528 Indels: 8
 Gaps: 4

US-09-622-964-3 (1-585) x AY064707 (1-1289)

QY 155 GlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisGlnLeuGluLysLeu 174
 DB 6 CAACACCTGGTGAAGACGAGCTTTATGACCCCTCGAACAACAGACTTGAACAAACG 65
 QY 175 SerLeuProHisAsnMetPheThrValProTyrValThrPheAlaAsnLeuSerMetLys 194
 DB 66 AGCTTGCCACACAACTCATCTCGATGCCCTGGGTGGTGGTCCAACTGTCAACGAAG 125
 QY 195 AlaTrpLeuGlyGlyArgIleArgAspProIleLeuGlnSerLeuLeuAsnGlnMet 214
 DB 126 GCGTGGATTGGGGGTGCAATCCGGGACCCCTGCTCCACAGACTGTGTCGCAAGATG 185
 QY 215 AsnThrLeuArgThrGlnCysGlyHisIleuTyrAlaTyrAspTrpIleSerIleProLeu 234
 DB 186 AACACCTTGCGTACTCACTGTGGACACCTGTATGCTTACGACTGATCATGTCGCGTG 245
 QY 235 ValTyrThrGlnValIleThrValAlaValTyrSerPhePheLeuThrCysLeuValGly 254
 DB 246 GTGTACACTCAGGTGGTGGACCGGTGATACAGCTTCTTCTGCGCTTGTCTGGTGGG 305
 QY 255 ArgGlnPheLeuAsnProAlaLysAlaTyrProGlnHisGlnLeuAspLeuValPro 274
 DB 306 AGGCAAGTTTCTGAACCCAGCCAGCCCTACCCCGGCATGATGAGACCTGTTGTACC 365
 QY 275 ValPheThrPheLeuGlnPhePhePhePheValGlyTyrPheLysValAlaGlnLeu 294
 DB 366 CTGTTCACGTTCTCGACGCTTCTTCTATGCGCGGTGGTGAAGGTGGCAGACACGCTC 425
 QY 295 IleAsnProPheGlyGluAspAspAspPheGluThrAsnTrpIleValAspArgAsn 314
 DB 426 ATCAACCCGTTTGGAGAGATGATGACAGCTTTGAGACCACTGGATTTGACAGAGAGC 485
 QY 315 LeuGlnValSerLeuLeuAlaValAspGlnMetHisGlnAspLeuProArgMetGluPro 334
 DB 486 TTTCAGAGTGTCCCTGTGGCTGTAGATGATGACACACGACCTGCACCCATGGAGCGA 545
 QY 335 AspMetTyrTrpAsnLysProGluProGlnProTyrThrAlaAlaSerAlaGlnPhe 354

DB 546 CATATGTACTGGAAATGATCCAGAACACATCCCCCTACAGAGCTGTTCGCCACGCT 605
 QY 355 ArgAlaGlnSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGlnLeuMetLys 374
 DB 606 CGCCGACCCCTCTTTTGGGCTCCACCTTCAATATCACTGTGGTGAAGAGACACTGGAG 665
 QY 375 PheGlnProAsnGlnGluAspGlnGluAspAlaHisAlaGlyIleIleGlyArgPheLeu 394
 DB 666 TTCCAGCGC-----GAGGAGAGAGAGAGCGCACACCGCATCTTGGCCACTTCTGT 719
 QY 395 GlysLeuGlnSerHisAspHisHisProProArgAlaAsnSerArgThrLysLeuLeuTrp 414
 DB 720 GGGGTGCATTCACGACGACACACACCCCAAGCAAACTCAAGAGCAAACTCTGTGG 779
 QY 415 ProLysArgLysSerLeuLeuHisGlnGlyLeuProLysAsnHisLysAlaValGln 434
 DB 780 CCCAAGAAAGAGAGCCATTTCACAGAGGCGCACCCCAAGAACCTCAGGGGGCCAGCTG 839
 QY 435 AsnValArgGlyGlnGluAspAsnLysAlaTrpLysLeuLysAlaValAlaPheLys 454
 DB 840 GACTCTAGCGACCAAGAGAGACAGACAGCCCTGGAGG-----GAGGGTGGCTTCAG 890
 QY 455 SerGlyProLeuTyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerPro 474
 DB 891 TCCGCTGGCCCTGTGTGGAGAGCGGGGTACACAGCGCCCAACAGACACCCCTGGCCAC 950
 QY 475 ThrProMetPhePheProLeuGluProSerAlaProSerLysLysSerValThrGly 494
 DB 951 ACCCTATGATGTCTCCACCCGAAAGAGTACAGCCCTTATGAGTTCGACAGAGTCTGGG 1010
 QY 495 IleAsp-----ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysSer 512
 DB 1011 ATAGATGAGCGCTGCCAAGACCAAGAGCTTACCGCTGCCATCCACATCAAGAGAGT 1070
 QY 513 PheGlnLeuLeuSerGlySerAspGlyAlaLeuMetGlnHisProGluValSerGlnVal 532
 DB 1071 TTTCAGTCTCTCCAGAGAGCGCCGAGCGCTGGGGGACCCCTCGAAGGAGTCAAGT 1130
 QY 533 ArgArgLysThrValGlnPheAsnLeuThrAspMetProGluIleProGluAsnHisLeu 552
 DB 1131 AGGAGGAAAGCGGTGAGTAACTGCGCGATCTGTCTGGAGCGCCCGCA---CATCTC 1187
 QY 553 LysGluProLeuGlnGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAsp 572
 DB 1188 AAAGAACCGAATTGTGGACCCACCATGGCGATACACGCTATATCTCAAGAACACCGGAT 1247
 QY 573 ProTyrTrpAlaLeuGluAsnArgAspGluAlaHisSer 585
 DB 1248 CCTACTGTGGCTTGGAAACAGAGGATGAACACACTCC 1286

RESULT 5
 AC003025 112309 bp DNA linear PRI 23-JUL-1998
 LOCUS Human Chromosome 11p12.2 PAC clone pD466a11, complete sequence.
 AC003025
 DEFINITION AC003025.1 GI:3337308
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 112309)
 Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basile, M.,
 Buettner, J., Bumeister, R., Card, P., deaillboat, F., Dunn, J.,
 English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G.,
 Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, R., Loo, K.N.,
 Major, T., McFarland, D., Newton, J., Osborne-Lawrence, S.,
 Schageman, D., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
 HUGS Submission
 Unpublished
 2 (bases 1 to 112309)
 Evans, G.A., Athanasiou, M., Braddury, P., Brignac, S., Bumeister, R.,

QY 21 LeuLeuCystrpArgglySerIleTyrLysLeuLeuTyrGluPheLeuIlePheLeu 40
 DB 165 CTGCTGTCGGGGGGGACACATCTACACCTGCTATATGGCGAGTCTTAACCTTCCTG 224
 QY 41 LeuCystrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGluGluGlnLeu 60
 DB 225 CTCTGCTACATCATCTCCCTTTATTATAGCTGGCCCTCAAGGAAACAAACACTG 284
 QY 61 MetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
 DB 285 ATTTTGAAGAACTGACTCTGTAATTCGACAGCTACATCCAGCTCATCTCCATTTCTTC 344
 QY 81 ValLeuGlyPheTyrValThrLeuValValThrArgTyrPheAsnGlnTyrGluAsnLeu 100
 DB 345 GTGCTGGCTTCTACGTAGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 404
 QY 101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluLysAspGluGln 120
 DB 405 CCGTGCCCGACCGCTCATGAGCTGTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 464
 QY 121 SerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuLysValLeuIleLeu 140
 DB 465 GCGCGCTCTCTGGGGGACAGGCTCATCTCCAGCTGAGCTGGCTGGCTGGCTGGCTGG 524
 QY 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
 DB 525 CCGAGGCTCAGCAGCCAGCTACAAAGGCTTCCCGAGCGCCAGCAGCTGGTGGCAAGA 584
 QY 161 GlyPheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet 180
 DB 585 GCGTTTATGACTCCGCGCAACACAAAGGTTGGAAACTAGCCTTACACACAAACATG 644
 QY 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyGlyArg 200
 DB 645 TTCTGGGTGCTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 704
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
 DB 705 ATCCCGGAGCCCTATCTCTGCTCCAGAGCCGTGAACGAGATGACACCTTGGTACTCG 764
 QY 221 CysGlyHisLeuTyrAlaTyrAspThrPheIleSerIleProLeuValTyrThrGlnVal 240
 DB 765 TGTGACACCTGTAATGCTTACAGCTGATATATCCACTGCTGTATACACAGGTGGTGG 824
 QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 DB 825 ACTGAGGGGTGTAAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 884
 QY 261 AlaLysAlaTyrProGlyHisGluLeuAspLeuValProValPheThrPheLeuGln 280
 DB 885 GCCAAGGCTTACCTGCGGCAATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
 QY 281 PhePhePheTyrValGlyTyrLeuLysVal 290
 DB 945 TTCTTCTTCTATGTTGGTGGCTGAAGGT-GGGCTCTCCAGAGCCCTGCTGGCTGGAG 1003
 QY 290 290
 DB 1004 GCATGGCCAGAGGGGTATGCGCCAGACACTGCTTGAAGCAGAGATGCTGACAGAAAG 1063
 QY 290 290
 DB 1064 GAAGGTCTCAGGGGTAGAAACAGCAGCAGCGTGTGGCCACACCTGTAATCCAGCTAC 1123
 QY 291 291
 DB 1124 TCGGAGAGCTCAGCAGAGAAATGCTTGAACCCGGGAGGGGAGGTTGTGGTGGCAAG 1183
 QY 293 GlnLeuIleAsnProPheGlyGluAspAspAspPheGluThrAsnTyrIleValAsp 312
 DB 1184 CAGCTCATCAACCCCTTGGAGAGATGATGATTTTGGACCACTGATTTGTCTGCAC 1243
 QY 313 ArgAsnLeuGlnValSerLeuLeuAlaValAspIleMetHisGlnAspLeuProArgMet 332

DB 1244 AGGAATTTGAGAGTGTCCCTGTTGGCTGTGGATGAGATGACACAGGACCTGCGGATG 1303
 QY 333 GluProAspMetTyrTrpAsnLysProGluProGluProProTyrThrAlaIleSerAla 352
 DB 1304 GAGCCGGACATGTAAGTGAATTAAGCCCGAGCCACAGCCCTTACACAGCTGCTCCGCC 1363
 QY 353 GlnPheArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGluGlu 372
 DB 1364 CAGTCCGTGAGACCTTCCTTATGGGCTCCACCTCAACATCAAGCTGAGCAAAAGAGAG 1423
 QY 373 MetGluPheGlnProAsnGlnLysAspGluGluAspAlaHisAlaGlyIleIleGlyArg 392
 DB 1424 ATGAGGTTCCAGCCCAATAGAGGAGCAGAGAGAGATGCTCACCTGTCATCATTTGGCCGC 1483
 QY 393 PheLeuGlyLeuGlnSerHisAspHisProProArgAlaAsnSerArgThrLysLeu 412
 DB 1484 TTGCTAGGCTGAGTCCCATATCATCCATCTCCAGGCGAACTCAAGGACCAACTA 1543
 QY 413 LeuTyrProLysArgGluSerLeuLeuHisGluGlyLeuProLysAsnHisLysAlaAla 432
 DB 1544 CTGTGGCCCAAGAGGAGATCCCTTCTCCAGAGGCGCTGCCCAAAACCAAGCAGCC 1603
 QY 433 LysGlnAsnValArgGlyGlnLysAspAsnLysAlaTyrLysLeuLysAlaValAspAla 452
 DB 1604 AAACAGAACGTTAAGGGCCAGAGAACAAACAAAGGCTTGAAGCTTGAAGCTGTGACGCC 1663
 QY 453 PheLysSerGlyProLeuTyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeu 472
 DB 1664 TTCAAGTGTGCCCACTGATATCAGAGGCCAGCTTACAGTGGCCCAAGACGCCCTTC 1723
 QY 473 SerProThrProMetPhePheProLeuGluProSerAlaProSerLysLeuHisSerVal 492
 DB 1724 ACCCCACCTCCCATGTTCTTCCCTAGAACCATCAGCCCGCTCAAAAGCTTCACAGTGC 1783
 QY 493 ThrGlyIleAspThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSer 512
 DB 1784 ACAGGCAATGACACCAACAAACAAAGCTTAAAGACTGTGATGCTTGGGCCCAAGAAAGT 1843
 QY 513 PheGluLeuLeuSerGluSerAspAlaLysMetGluHisProGluValSerGlnVal 532
 DB 1844 TTTGAATTCCTCTCAGAGAGCATGGGGCTTATGAGGACCCAGAAAGTATCTCAAGTG 1903
 QY 533 ArgArgLysThrValGluPheAsnLeuThrAspMetProGluIleProGluAsnHisLeu 552
 DB 1904 AGGAGAAACCTGTGAGATTAACTGACGAGTATGCCAGAGATCCCGGAAATCATCACTTC 1963
 QY 553 LysGluProLeuGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAsp 572
 DB 1964 AAAGAACCTTTGGAACAAATACCAACCAACATACACTACACTCAAGATCAATGATGAT 2023
 QY 573 ProTyrTrpAlaLeuGluAsnArgAspGluAlaHisSer 595
 DB 2024 CCTTATTTGGCTTGGAAACAGGAGTGAAGACATTC 2062
 RESULT 4
 AY064707 1289 bp mRNA linear MAN 19-AUG-2002
 LOCUS Sus scrofa bestrophin mRNA, partial cds.
 DEFINITION
 ACCESSION AY064707
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 Marmorstein,L.Y., McLaughlin,P.J., Stanton,J.B., Yan,L., Crabb,J.W.
 Bestrophin Interacts Physically and Functionally with Protein
 Phosphatase 2A
 JOURNAL J. Biol. Chem. 277 (34), 30591-30597 (2002)
 PUBMED 12058047

```

Db      705 ATCCGGGACCTTATCTGCTCCAGAGCTGCTGAACGATGAAACACTTCGTAATCAG 764
Qy      221 CysGlyHisLeuTyrrAlaTyrrAspTrpIleSerIleProLeuValTyrThGlnValAl 240
Db      765 TGTGGACACCTGATGCTGCTACGACTGATGATTAATCCACAGGGGTATACACAGCTGGTG 824
Qy      241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db      825 ACTGTGGCGGTGTACAGCTTCTTCCTGACTGTCTACTTGGGCGGCGATTCTGACACCA 884
Qy      261 AlAlaValAlaTyrProGlyHisGlnLeuAspLeuValAlProValPheThrPheLeuGln 280
Db      885 GCCAAGGCTTACCTGGCCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
Qy      281 PhePhePheTyrValGlyTyrPheLeuValAlaGlnGlnLeuLeuAsnProPheGlyGlu 300
Db      945 TTCTCTCTTATCTGTGGCTGCTGGAAGTGGCAGAGCTCATCAACCTCTTGGAGAG 1004
Qy      301 AspAspAspAspPheGluThrAsnTrpIleValAlaAspArgAsnLeuGlnValSerLeuLeu 320
Db      1005 GATGATGATGATTTTGAACCAACTGATGTGACAGAAATTTGACAGTGTCTCTCTG 1064
Qy      321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLys 340
Db      1065 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124
Qy      341 ProGluProGluProProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db      1125 CCCGAGCCACAGCCCCCTACACAGCTGCTCCGCCCATCTCCGTGAGAGCTCTCTTATG 1184
Qy      361 GlySerThrPheAsnIleSerLeuAsnLysGlnGluMetGlnPheGlnProAsnGlnGlu 380
Db      1185 GGTCTCACCTTCAACATCAAGCTGTAACAAGAGAGATGAGATGATCCAGCCCATCGAGAG 1244
Qy      381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
Db      1245 GACGAGGAGGATGCTCAGCTGCTGATCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1304
Qy      401 HisHisProProArgAlaAsnSerArgThrLysLeuLeuTrpProLysArgGlnSerLeu 420
Db      1305 CACCATCTCCCGCAGGCAAACTCAAGCAAACTCACTGTGCCCCCAAGAGGGAATCTCTT 1364
Qy      421 LeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgLysGlnGlu 440
Db      1365 CTCACGAGGGGCTGCCCCCAAAACCAAGCAGCAGCAAAACAGCTTAAAGGGCCAGAGA 1424
Qy      441 AspAsnLysAlaTyrPheLysLeuValAlaValAspAlaPheLysSerGlyProLeuTyrGln 460
Db      1425 GACAACAGGCTGGAACCTTAAGCTGTGGAGCGCTTCAAGTCTGCCCACTGATCAG 1484
Qy      461 ArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPrometPhePhePro 480
Db      1485 AGGCGAGGCTACTACAGTGGCCCCACAGAGCGCCCTCAGACCCCATCTCTCTCTCC 1544
Qy      481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
Db      1545 CTAAACCATCAGGCGCTCAAAAGCTTCAAGTCTCAGAGCAGCATAGACCAAAAGACAA 1604
Qy      501 SerLeuLysThrValSerSerGlyAlaLysSerPheGluLeuLeuSerGlnSerAsp 520
Db      1605 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAAATGTCTCTCAGAGAGCAT 1664
Qy      521 GlyAlaLeuMetGlnHisProGluValSerGlnValAlaArgArgLysThrValGluPheAsn 540
Db      1665 GGGGCGCTTGTGAGCAGCAGCAAGAGTATCTCAAGTGAAGGAGAAATCTGTGAGTTAAC 1724
Qy      541 LeuThrAspMetProGluLysLeuProGluLysAsnHisLysLeuGluProLeuGlnGlnSerPro 560
Db      1725 CTGACGATATGCGACAGAGATCCCGAAATACCTTCAAGAACCTTGGAAACATCACCA 1784
Qy      561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTrpAlaLeuGluAsnArg 580

```

```

Db      1785 ACCAACATACACACTACACTCAAGATGATGATCTTATGTGGCCTTGAAACAGG 1844
Qy      581 AspGluAlaHisSer 585
Db      1845 GATACACACATCTCC 1859

RESULT 3
AF057170
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITILE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
JOURNAL
FEATURES
source
gene
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/1..2420
/gene="VMD2"
/105..1412
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
splice isoform; contains exon 7"
/product="bestrophin"
/protein_id="AAC64344.1"
/db_xref="GI:3335161"
/translation="MTITITQVQVNAARLGSFRLLCWRGSIYKLYGEPFLICY
IIRITRLATEEQDMEKLLRLCDSTYQIPIISFVAGFVTLVTRMNOYELPA
PDRMLSVSGVEBEKDEGRLLRLRLAYANGLVILRSYTAIVKRPSPALNOY
GFMPDABEKOLEKSLPHNMFVPMVPMFASNMALGRIKRDILLOSILNENYTL
TOCGHLYADWISIPLYTOVTVAVYSFELICVGRQFLPAKAYRHELDVYVPF
TFLEFFYVGLKXGSLRALIGMHGGRGHOQLERMOCKEKRVSVESSQAMWRT
PVIRATREANGESLEBERRLMMOSSSTFLERMMILRPTGLSTGICRCPCWLMMR
CTRCLEHSRCTGTSISHSPTQLLPSSVPEPLMAPPSTSA"
BASE COUNT 605 a 693 c 612 g 508 t
ORIGIN

Alignment Scores:
Pred. No.: 2,32e-256 Length: 2420
Score: 3041.50 Matches: 582
Percent Similarity: 89.13% Conservative: 0
Best Local Similarity: 89.13% Mismatches: 3
Query Match: 97.48% Indels: 69
Gaps: 1

US-09-622-964-3 (1-585) x AF057170 (1-2420)
Qy      1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db      105 ATGACCATCACTTACACAGCAGGATGAGTAAATGCCCGCTTAAAGCTCTCTCCGCGCTG 164

```

```

Db      1289 CACCATCTCCAGGCGCAACTCAGACCAAACTCTGTGGCCCAAGAGGGATCCCTT 1348
Qy      421 LeuHISGLUGlyLeuProLysasnHISLysAlaAlaLysGlnAsnValArgLysGlnGlu 440
Db      1349 CTCACAGAGGCGCTGCCCAAAACCAAGAGCAGCAACAGAACTGAGGGCCAGGA 1408
Qy      441 AspaNtLysAlaTrpLysLeuLysAlaValAlaPheLysSerGlyProLeuTyrgln 460
Db      1409 GCAACAAAGCGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCTGCTGTATCAG 1468
Qy      461 ArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrProMetPhePro 480
Db      1469 AGGCAGAGCTACTACAGTGGCCACAGACACCCCTCAGCCCTCCATGTTCTTCCCC 1528
Qy      481 LeuGluProSerAlaProSerLysLeuHISerValThrGlyIleAspThrLysAspLys 500
Db      1529 CTGAAACCATCAGCCCGCTCAAAAGCTTCACAGTGTACAGCGCATAGACCAAGACAA 1588
Qy      501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerLysSerasp 520
Db      1589 AGCTTAAAGACTGTGAGTCTGTGGGCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGAT 1648
Qy      521 GlyAlaLeuMetGluHISProGluValSerGlnValArgArgLysThrValGluPheasn 540
Db      1649 GGGCGCTGTGATGAGCAGCACCAAGATATCTCAAGTGAAGAGAAACTGTGAGATTATAC 1708
Qy      541 LeuThrAspMetProGluIleProGluLysasnHISLysGluProLeuGlnGlnSerPro 560
Db      1709 CTGACGGATATGCGCAGAGATCCCGCAAAATCACCCTCAAGAAACCTTGTGAACAAATCACCA 1768
Qy      561 ThrAsnIleHISThrThrLeuLysAspHISMetAspProGlyTyrTrpAlaLeuGlnAsnArg 580
Db      1769 ACCAACAATACACTACACTACCTCAAGATACATGATGCTTATTTGGGCTTGAAAAACAGG 1828
Qy      581 AspGluAlaHISser 585
Db      1829 GATGAGACACATTC 1843

RESULT 2
AF057169 2210 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
DEFINITION AF057169
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2210)
AUTHORS Sandgren,O., Forsman,K., Holmgren,G., Andreasson,S., Vulic,M.,
Berger,A.A., McGarvey-Dugan,V., Figueroa,D., Austin,C.P.,
Metzger,M.L., Caskey,C.T. and Wadelius,C.
TITLE Identification of the gene responsible for Best macular dystrophy
JOURNAL Nat. Genet. 19 (3), 241-247 (1998)
MEDLINE 98324772
PUBMED 9662395
REFERENCE 2 (bases 1 to 2210)
AUTHORS Petrushkin,K.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA
FEATURES
source
1..2210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2210
/gene="VMD2"

```

```

CDS
105..1862
/gene="VMD2"
/Note="Best's macular dystrophy gene; BMD; alternatively
splice isoform"
/codon_start=1
/product="bestrophin"
/protein_id="AAC64343.1"
/db_xref="GI:3335159"
/translation="MTITTSQVANNARIGSFSRLILCWNGSITKLYGSEFLFLCY
IRFLYRATLAEQQLMEKLLTXDSTYQQLPISVLCFTYTLVTRMNDENLPM
DRLMSLYSGFEQDEQRLRLRLIRANLGNVILISVSTAVYKRPSSQHLVQA
GMPTRAEHQLEKLSLPHMEVWVWFWFANLSMKAMLGRIIDPLLOSILNEMTLR
TQCGHLVADWISIPLYVTVVAVYSEFLCYLGRQFLPAKAPGHEHLVVVF
TELQEFYVGMKVAEOLINPGEQDDDEITWYIDRNLOVSLAVDEHOLPRMEP
DMYANKPEPOPPTAASQFRASFSMGSFNLISLKEEERPNQDEDAAGITIG
FLGLOSHDHHPRRASRKLMPKRESLHELHELPRNKAQKNVQEDNKAMKLAV
DAFKAPLYKQRGYISAPOTLSPIMFEPFLPSAPSLHSVYGLDITDKSLKTVSSG
AKKSEFLSESDGLMEHPREVSOVRKRVFENLTDMPLEIPENHLEPLEOSPNTIHT
LKDHDPPYWALENRDEAHS"
BASE COUNT 556 a 647 c 531 g 476 t
ORIGIN

Alignment Scores:
Pred. No.: 2,08e-262 Length: 2210
Score: 3110.00 Matches: 583
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 2
Query Match: 99.68% Indels: 0
Df: 9 Gaps: 0

US-09-622-964-3 (1-585) x AF057169 (1-2210)
Qy 1 MetThrIleThrThrThrSerGlnValAlaAsnAlaArgLysGluSerPheSerArgLeu 20
Db 105 ATGACCATCATTACACAGCCCAAGTGGCTAATGCCCTTAGCTCTTCCCGCGCTG 164
Qy 21 LeuLeuGlySerPheGlySerIleThrLysLeuLeuTyrgLysPheLeuIlePheLeu 40
Db 165 CTGCTGTCTGCGCGGCGGAGCATCTACAAAGCTGCTATATGCGAGATTCTTATCTTCTG 224
Qy 41 LeuGlyTyrTrpIleIleArgPheIleThrArgLysAlaLeuThrGlnGlnGlnLeu 60
Db 225 CTCTGCTACTACATCATCCGCTTATATAGCGTGGCCCTCAGGAAACAAACAGCTG 284
Qy 61 MetPheGluLysLeuThrLeuTyrcysAspSerTyrIleGlnLeuIleProIleSerPhe 80
Db 285 ATGTTGAGAAACAGACTGTGTATGCGACAGCTACATCCAGCTCATCTCCATTCTTC 344
Qy 81 ValLeuGlyPheTyrgValThrLeuValValThrArgTyrPheAsnGlnTyrgLysAsnLeu 100
Db 345 GTGCTGGGCTTCTACGTGACCGCTGCTGTCACCGCTGCTGGAACCAATGACGAACTG 404
Qy 101 ProThrProAspArgLeuMetSerLeuValSerLysPheValGluGlyLysAspGluGln 120
Db 405 CCGAGGCCCAAGCCCTCATAGAGCTGTGCGGGCTTCGTCGAAGGCAAGAGAGACCA 464
Qy 405 CCGAGGCCCAAGCCCTCATAGAGCTGTGCGGGCTTCGTCGAAGGCAAGAGAGACCA 464
Db 121 SerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGluGlyAsnValLeuLeu 140
Qy 465 GCGCGGCTGCTGGCGGCGCACGCTCATCCGCTACGCCCAACCTGGGCAACGTGCTATCTG 524
Db 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHISLeuValGlnAla 160
Qy 525 CGCAGCGCTCAGCACCGGAGTCTCAAGCGCTTCCCGCCAGCCGACCTGTGTCAGACA 584
Db 161 GlyPheMetThrProAlaGluHISLysGlnLeuGlnLysSerLeuPheAsnMet 180
Qy 585 GCGTTATATGACTCCGCGAGAACACAGCACTTGAGAAACTGAGCTTACCAACACAAATG 644
Db 181 PheThrValProThrValTyrPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
Qy 645 TTCCTGGTGCCCTGGGCTGTGTTGCCAACCTGTCAATGAAGGCGTGTGAGAGTGA 704
Db 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:07:41 ; Search time 4477.8 Seconds
(without alignments)
3802.121 Million cell updates/sec

Title: US-09-622-964-3
Perfect score: 3120
Sequence: 1 MTITVTSQVANARLGSFSL.....TLKHMDPYALENDEAHS 585

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV-x1h
-Q/cgn_1/USPRO.spool/US09622964/tunat_22072003.101118.26013/app.query.fasta_1.2069
-DB-genemb1 -QMT-fastap -SOFFI-rge -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST-45
-DOCLIN-200 -THR SCORE-pct -THR MAX-100 -THR MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09622964.ecgn_1.1.7526.grunat_22072003.101118.26013 -NCPU-6 -ICPU-3
-NO.MMAP -IARGOUTDIR -NBS-SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : GenEmb1:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_in:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_pl:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3110	99.7	2171	9 AF073501	AF073501 Homo sapi
2	3110	99.7	2210	9 AF057169	AF057169 Homo sapi
3	3041.5	97.5	2420	9 AF057170	AF057170 Homo sapi
4	1701	54.5	1289	4 AY064707	AY064707 Sus scrofa
5	1655.5	53.1	112309	9 AC003025	AC003025 Human chr
6	1652.5	53.0	142092	9 AF139813	AF139813 Homo sapi
7	1652.5	53.0	196080	9 AC004758	AC004758 Homo sapi
8	1645.5	52.7	163915	9 AC087451	AC087451 Homo sapi
9	1645.5	52.7	166867	9 AP003733	AP003733 Homo sapi
10	1636.5	52.5	133683	2 AC084857	AC084857 Homo sapi
11	1433	45.9	160169	2 AC051664	AC051664 Homo sapi
12	1332	42.7	2500	9 AC096459	AC096459 Homo sapi
13	1296	41.5	1506	9 AF440758	AF440758 Homo sapi
14	1286.5	41.2	1956	10 BC031186	BC031186 Mus muscu
15	1286.5	41.2	1957	10 BC019528	BC019528 Mus muscu
16	1284	41.2	1908	9 AF440757	AF440757 Homo sapi
17	1209.5	38.8	2045	9 AC000139	AC000139 Homo sapi
18	1199	38.4	1861	9 HSWMD2P10	AF073499 Homo sapi
19	1142	36.6	706	9 AF218817	AF218817 Drosophila
20	1009	32.3	2874	3 AC019521	AC019521 Drosophila
21	1009	32.3	5610	3 AC019521	AC019521 Drosophila
22	851.5	27.3	32022	2 AC008139	AC008139 Drosophila
23	851.5	27.3	170675	3 AC009183	AC009183 Drosophila
24	851.5	27.3	192055	3 AC009183	AC009183 Drosophila
25	851.5	27.3	221888	3 AC009183	AC009183 Drosophila
26	828	26.5	163514	3 AC091227	AC091227 Drosophila
27	828	26.5	274351	3 AC003531	AC003531 Drosophila
28	825	26.4	125804	2 AC019853	AC019853 Drosophila
29	801	25.7	125804	2 AC019853	AC019853 Drosophila
30	801	25.7	163514	3 AC091227	AC091227 Drosophila
31	801	25.7	274351	3 AC003531	AC003531 Drosophila
32	713.5	22.9	147750	9 AC018761	AC018761 Homo sapi
33	708	22.7	90487	9 AL592166	AL592166 Human DNA
34	676.5	21.7	22973	3 AF025458	AF025458 Caenorhabd
35	640	20.5	181453	2 AC130166	AC130166 Rattus no
36	631	20.2	39590	3 CHR39K23	CHR39K23 Caenorhabd
37	624.5	20.0	219043	2 AC073812	AC073812 Mus muscu
38	608.5	19.5	330724	2 CEY67H2	CEY67H2 Caenorhabd
39	607.5	19.5	39333	3 CEB0564	CEB0564 Caenorhabd
40	605.5	19.4	22973	3 AF025458	AF025458 Caenorhabd
41	562.5	18.0	37977	3 L16621	L16621 Caenorhabd
42	549.5	17.6	27394	3 CEC29F4	CEC29F4 Caenorhabd
43	528.5	16.9	17402	3 CEC29F4	CEC29F4 Caenorhabd
44	525.5	16.8	2912	9 BC028087	BC028087 Homo sapi
45	509.5	16.3	105126	2 AC014828	AC014828 Drosophila

RESULT 1

ALIGNMENTS

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:10:16 ; Search time 2336.65 Seconds

(without alignments)
4054.679 Million cell updates/sec

Title: US-09-622-964-3
Perfect score: 3120
Sequence: 1 MTITVTSQVAVARLGSFSL.....TLKDHMPYALLENDEAHS 585

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US0962264/runat_22072003_101119_26026/app-query.fasta_1.2069
-DB=EST -OPMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pcpt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USFR=US0962264.ecgn.1.1.4188.grunat.22072003_101119_26026 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_luv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2672.5	85.7	2453	11	BC015220
2	1640	52.6	1699	11	AK006549
3	1389.5	44.5	963	14	BQ879880
4	1020	32.7	773	10	BE410951
5	1011	32.4	593	10	BE385296
6	951	30.5	735	13	BT756228
7	941	30.2	537	14	BM691456
8	935	30.0	930	14	BQ436824
9	934	28.9	666	13	B1480798
10	915	28.3	526	14	BM685501
11	830	26.6	793	13	BG961794
12	813	26.1	469	9	AA307119
13	807	25.9	455	12	BC013943
14	793	25.4	522	14	BM707649
15	781	25.0	531	14	BM685396
16	779	25.0	508	10	BE236901
17	761.5	24.4	601	13	BG951790
18	761.5	24.3	654	10	BE679896
19	758	24.3	566	14	BM718338
20	744.5	23.9	766	10	BE275846
21	735.5	23.6	484	14	N31453
22	693	22.2	613	9	AL654606
23	668	21.4	1067	13	BM562042
24	667.5	21.4	722	9	AL627604
25	663.5	21.3	566	13	B1443895
26	658	21.1	371	14	BM694530
27	656	21.0	628	13	BU069061
28	643	20.6	732	13	BG965806
29	637.5	20.4	603	9	AA205892
30	608	19.5	657	14	BM397157
31	592.5	19.0	672	13	BM576476
32	574.5	18.4	1111	17	CNS04DPM
33	574	18.4	514	13	BT343182
34	572	18.3	577	10	BE189780
35	570.5	18.3	514	13	BE669309
36	566	18.1	966	12	BF582222
37	563.5	18.1	599	13	BG951382
38	561	18.0	635	14	BO396082
39	544	17.4	649	10	BE189780
40	543	17.4	503	14	BM685122
41	543	17.4	508	14	BM932117
42	533	17.1	525	14	BO345562
43	529	17.0	605	14	BO391968
44	528.5	16.9	730	13	BG963329
45	527	16.9	585	14	BM707948

ALIGNMENTS

RESULT 1
LOCUS BC015220 2453 bp mRNA
DEFINITION Homo sapiens, clone IMAGE:3877806, mRNA.
ACCESSION BC015220
VERSION BC015220.1 GI:21955361
KEYWORDS
SOURCE HTC.
ORGANISM Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 2453)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.simg.stanford.edu>
 Contact: (Dickson, Mark) mcdepaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRK Plate: 14 Row: b Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

This clone has the following problem: Incomplete processing.

FEATURES

source

Location/Qualifiers
 1. 2453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3877806"
 /tissue_type="lung, large cell carcinoma"
 /clone_id="NH_MGC_68"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 BASE COUNT 633 a 711 c 606 g 503 t
 ORIGIN

Alignment Scores:

Prod. No.: 1.35e-294 Length: 2453
 Score: 2672.50 Matches: 505
 Percent Similarity: 94.57% Conservative: 0
 Best Local Similarity: 94.57% Mismatches: 2
 Query Match: 85.66% Indels: 27
 DB: 11 Gaps: 1

US-09-622-964-3 (1-585) x BC015220 (1-2453)

QY 52 LeuAlaLeuThrgluGluGlnGlnLeuMetPheGluLysLeuThrLeuTyrCysAspSer 71
 DB 558 CTGGCCCTCAGCGAAGAACAGACGCTGTTGAGAACTGCTGTATTCGACAGC 617
 QY 72 TTTTleGlnLeuThrProIleSerPheValLeuGlyPheTyrValThrLeuValValThr 91
 DB 618 TACATCCAGCTCATCCCAATTCCTTCGCTGGCTTCACTGACGCTGCTGAGAC 677
 QY 92 ATGTTTPTAsnGlnTyrGlnAsnLeuProTrrProAspArgLeuMetSerLeuValSer 111
 DB 678 CCGTGTGAGAACAGTACGAGACCTGCGTGGCCGACCGCTCTAGAGCGGTGTGG 737
 QY 112 GlyPheValGluGlyLysAspGlnGlnSerArgLeuLeuArgThrLeuIleArgTyr 131
 DB 738 GCGTTCGTGAGAGGAGGAGGAGGAGCGCCGCTCTGCGGCGACGCTCATCCGCTAC 797
 QY 132 AlaAsnLeuGlyAsnValLeuIleLeuArgSerValSerThrAlaValTyrIleAspPhe 151
 DB 798 GCCAACCTGGGACAGCTGCTCATCTCCGACGCTGACGACGACGCTCATCCGCTTC 857
 QY 152 ProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeu 171
 DB 858 CCCAGCGCCACACCTGCTGCAAGGAGCTTATGACTCCGACAGAACACAGCAGTTGG 917
 QY 172 GtLysLeuSerLeuProHisAsnMetPheTrrValProTrrValTrrPheAlaAsnLeu 191
 DB 918 GAGAAAGTGAAGCTGACACACAAACATGTTCTGGGTGCGCTGGGTGGTTTGCACAACTCG 977

QY 192 SerMetLysAlaTrrPheGluGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeu 211
 DB 978 TCAATGAGAGGCGTGGTGGAGGTCGAAATCCGGGACCTATCTGCTCCAGACCTCTGTG 1037
 QY 212 AsnGlnMetAsnThrLeuArgThrGlnCysGlnHisLeuTyrAlaTyrAspTrrIleSer 231
 DB 1038 AACAGAGTGAACACCTTCCGTCACAGTGTGACACCTGTATGCTTACGACTGATTAAGT 1097
 QY 232 IleProLeuValTyrThrGlnValValThrValAlaValTyrSerPhePheLeuThrCys 251
 DB 1098 ATCCCACTGGGTGTATACACAGGTGTGCTGCGGTGTACACTTCTCTCTGCTGCTGT 1157
 QY 252 LeuValGlyArgGlnPheLeuAsnProAlaLysAlaTyrProGlnHisGlnLeuAspLeu 271
 DB 1158 CTAGTGTGGGCGAGTTTCTGAAACCCAGCCCAAGGCTTACCTGGCCATGAGACTGAGACTTC 1217
 QY 272 ValValProValPheThrPheLeuGlnPhePheThrTyrValGlyTrrPheLysValAla 291
 DB 1218 GTTGTGCGCGTCTTACAGTTCCTGCACTTCTCTTATGTTGGCTGCTGAG----- 1271
 QY 292 GtLysLeuIleAsnProPheGlyGluAspAspAspPheGlnThrAsnTrrPleVal 311
 DB 1271 ----- 1271
 QY 312 AspArgAsnLeuGlnValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProArg 331
 DB 1272 -----GTGTCTCTGTGGCTGTGTGATGAGATGACACAGGACCTGCTCGG 1316
 QY 332 MetGluProAspMetTyrTrrPAsnLysProGluProGlnProTrrTrrAlaAlaSer 351
 DB 1317 ATGAGAGCGGACATGTACTGTGAATTAAGCCGACCCACAGCCCTTACACAGCTGCTTC 1376
 QY 352 AlaGlnPheArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGlu 371
 DB 1377 GCCAGTTCGCTGAGAGCTCTTATGAGGCTCCACCTTCAACATCACCTGAGCAACAAAG 1436
 QY 372 GtMetGluPheGlnProAsnGlnGluAspGlnGluAspAlaHisAlaGlyIleGly 391
 DB 1437 GAGATGAGATTCACAGCCCAATCAGAGAGAGAGAGATCTTACGCTGGCATCATTTGGC 1496
 QY 392 ArgPheLeuGlyLeuGlnSerHisAspHisHisProProArgAlaAsnSerArgThrLys 411
 DB 1497 CGCTTCCTAGGCTCGAGTCGATGCCATGATCACATTCCTCCAGGAGCAACCTGACAGCA 1556
 QY 412 LeuLeuTrrProLysArgGlnSerLeuLeuHisGluGlyLeuProLysAsnHisLysAla 431
 DB 1557 CTACTGTGGCCCAAGAGGAGATCCCTTCTCCAGAGGAGGCTGGCCAAAACCCACAGGCA 1616
 QY 432 AlaLysGlnAsnValArgGlyGlnGluAspAsnLysAlaTrrPysLeuLysAlaAsp 451
 DB 1617 GCCAAACAGAGCTTGGGGGCCAGAGACACAAAGCCCTGGAAGCTTAAGCTGTGGAC 1676
 QY 452 AlaPheLysSerGlyProLeuTyrGlnArgProGlyTyrTrrSerAlaProGlnThrPro 471
 DB 1677 GCTTCATAGTCTGCCCACTGATATCAGAGGCCAGCTACTACAGTCCCCACAGACACCC 1736
 QY 472 LeuSerProThrProMetPhePhePheProLeuGlnProSerAlaProSerLysLeuHisSer 491
 DB 1737 CTGAGCCCACTGCCATGTTTCTTCCCTTGAACCACTTACAGCCCGCTCAAGGCTTCAAGT 1796
 QY 492 ValThrGlyIleAspThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLys 511
 DB 1797 GTACAGGCAATGACACCAACAAACAAACAAACCTTAAGCTGTGAGTCTGGGCCCAAGAA 1856
 QY 512 SerPheGlnLeuLeuSerGlnSerLysArgLysAlaLeuMetGlnHisProGluValSerGln 531
 DB 1857 AGTTTGAATTCCTCTCAG 1916
 QY 532 ValArgArgLysThrValGluPheAsnLeuThrAspMetProGluIleProGlnAsnHis 551
 DB 1917 GTACAGAGAGAAAGCTGTGAGATTAACCTGAGAGAGATATCCAGAGATATCCCAAAATATCC 1976
 QY 552 LeuLysGluProLeuGlnSerProThrAsnIleHisThrThrLeuLysAspHisMet 571

Oy	77	ProIleSerPheValLeuGlyPheTyrValIhLeuValIAlhTyrTrpAsnIn	96
Db	61	CCATATCTCTCGTTCGGTTTCTATTTACATGGTGGTACAGCCGTGGTGGACAG	120
Oy	97	TyGluAsnLeuProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGly	111
Db	121	TACAGAACTTGGCCGTGGCCGACCGCTCATGATCCAGGTGTCTAGCTGTGTGAAGGC	180
Oy	117	LysAspGluGlnSerArgLeuArgLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlyAsn	136
Db	181	AAGGATAGGAAGGCCGTTTGCTCGCGCGACCGCTCATCCGCTACGCCATCTGGGCCAA	240
Oy	137	ValLeuIleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHis	156
Db	241	GTGGTATCTCGGCCACATCAGCACTCGGTGTACAAAGCCTTTCACACTTTCACAC	300
Oy	157	LeuValGlnAlaGlyPheMetThrProAlaGluHisLysGlnLeuGlyLysLeuSerLeu	176
Db	301	CTGGTGTACAGGTTTATATACCCATGGGGAACATAGCAGTTCGAGAAGTTGGCCTA	360
Oy	177	ProHisAsnMetPheTrpValProTrpValThrPheAlaAsnLeuSerMetLysAlaTrp	196
Db	361	CCACA-AACACATTTCTGGGTGCCCTGGGTGTGTGTTTCCAACTGTCAAGAAGCCCTAT	420
Oy	197	LeuGlyGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThr	216
Db	420	CTTGGAGTGGTAATCCGGGACACCGTCTGCTCCAGAGCCTGATGATGAGTGTGTACT	479
Oy	217	LeuArgThrGlnGlyGlnHisLeuTyrAlaTyrLysTrpIleSerIleProLeuValTyr	236
Db	480	TTCGATCTCAGTGTGTGACAGCTGTATGCTTACAGTACGATGAATGATCCATTTGGTGTAC	539
Oy	237	ThrGluValValThrValAlaValTyrSerPhePheLeuThrCysLeuValIglTyrGln	256
Db	540	ACACAGGTGGTGCACATGGCAGATAGAGCTTTTCTCTTCATCAGTTCGAGGAGGACAG	599
Oy	257	PheLeuAsnProAlaLysAlaTyrProGluGlnHisGlyLeuAspLeuValValProValPhe	276
Db	600	TTTCTGAACCCAAACAAGACTACCCAGGCGCAAGATGATGTGGTTGGTCTGCTCTTC	659
Oy	277	ThrPheLeuGlnPhePheThrTyrValIglTyrLeuLysValAlaGluGlnLeuIleAsn	296
Db	660	ACATCTCTCAATTTCTTATTTACATGGGCTGGCTCAAGCTGGCACAACCTCATCAAC	719
Oy	297	ProPheGlyGlnAspAspAspAspPheGluThrAsnTrpIleValAspArgLeuGln	316
Db	720	CCCTTCGGGAGAGCATATGATATTTTGACACTAACTGGATCATGACAGAAACCTGCAG	779
Oy	317	ValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProArgLeuGluProAspMet	336
Db	780	GTGTCCCTGTGTCCGTGGATGGATGCACCGAACTTCCCTCCCTGAAGCTGCATAG	839
Oy	337	TyTrpAsnLysPheGluProGlnProTrpTyrThrAlaAlaSerAlaGlnPheArgTyr	356
Db	840	TACTGGAAAGAGGACGCGCTCAGCGCCCTACACAGCTGTCTGCGCATCTGCGCGG	899
Oy	357	AlaSerPheMetGlySerThrPheAsnIleSerIleAsnLysGluGluMetLurPheGln	376
Db	900	CATTCTTTCAATGGGCTCCACCTTCACATCAACCCCTTAAGAAAGAAAGACTTAAGAGTTTGG	959
Oy	377	ProAsnGlnGlu-----AspGluGlnAspAlaHisAlaGlyLleIleGlyArg	392
Db	960	TCAAAAGAGGAGGTGACACGATTAAGAAAGAGTGGGTATAGACGACACCTATAGGCTGC	1019
Oy	393	PheLeuGlyLeuGlnSerHisAspHisIshIshProProlArgAlaAsnSerArgThrLysLeu	412
Db	1020	TTCTTAGACACTGCACAAACCAAAACATCACTTCCCTGAAGAGCTTAAAGCCAAACTA	1079
Oy	413	LeuTrpProLysArgGluSerLeuThrLeuHisGluGlyLeuProLysArgHisLysAlaAla	432
Db	1080	TTTGTTCTTAGAACCCCTCTCTC-----GAGGCGACGTGTAAAGAT-----	1129

Oy	433	LysGlnAsnValArgGlyGlnGluAspAsnLysAlaIleThrPheLysAlaValAlaAspAla	452
Db	1122	---GCCAACAGAAAAACCAAGAAAT-----GTCGTGAAATTTAAAGGCTTGACCTTC	1177
Oy	435	PheLysSerGlyProLeuTyGlnAArgProGlyTYrTYrSerAlaProGlnThrProLeu	472
Db	1173	TTGAATATGTTCCAAAGGTTTAAAGAGAGAGCTCCATTTGGGCCACAGGCACACAGC	1233
Oy	473	SerProThrPrometPhePheProGluProSerAlaProSerLysLeuHisSerVal	492
Db	1233	AGC-----CACCTTACTAGACAGTACAGACCTCCAGT-----TCACAC	1273
Oy	493	ThrGlyIleAspThrLysAspLysSerLeuysThrValSerSerGlyAlaLysLysSer	512
Db	1272	ACAGGT-----	1277
Oy	513	PheGluLeuLeuSerGluSerAspGlyAlaLeuMetGlnHisProGluValSerGlnVal	532
Db	1278	-----GANGGGGCTTCCACAGATTTCCAAAGAAATCTGTCACATG	1316
Oy	533	ArgArgLysThrValGluPheAsnLeuThrAspMetProGluIleProGluAsnHisLeu	552
Db	1317	AAAAAGAAATCTGGAGACTTTAACTTG---AACATTCACAGAGAGGCCACACAGAAACATCTT	1373
Oy	553	LysGlu---ProLeuGluGlnSerProThrAsnIleHisThrThrLysAspHisMet	571
Db	1374	CACACAGGCGCCTTGGACACAGATGTCACCAACAATATACAGGCTTAATGACAGAGATGCA	1433
Oy	572	AspProTYrThrAlaLeuGluAsnAspAla	583
Db	1434	GAGTCCAT-----CCCTACAGGATGTAAAGCT	1460

[illegible]

```

FEATURES
source
Location/Qualifiers
1..963
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6180359"
/_clone_id="Lupsk1_dorsal_root_ganglion"
/_sex="male"
/_tissue_type="dorsal root ganglia"
/_dev_stage="adult, 36 yr"
/_lab_host="DH10B"
/_note="Vector: PCMV-SpOTr6 (Life Technologies); Site:1;
Note1: Site:2: SalI, cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCACGCGTCCG-3' and

```

5' -GACTAGTTCTAGATCCGAGCGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies.

BASE COUNT 205 a 289 c 265 g 202 t 2 others

Alignment Scores:

Pred. No.: 4.21e-148 Length: 963
Score: 1389.50 Matches: 275
Percent Similarity: 89.598 Conserved: 9
Best Local Similarity: 86.758 Mismatches: 23
Query Match: 44.548 Indels: 10
DB: 14 Gaps: 5

US-09-622-964-3 (1-585) x B0879880 (1-963)

```

Oy 110 ValSerGlyPheValGlyLysAspGluInsSerArgLeuLeuArgThrLeuIle 129
Db 16 ATTTCCGGG---ATCGAAGCGAAGAGAGAGAGCGCGCTCTCCGCGCAGCTCAGC 72
Oy 130 ArgTyrAlaAsnLeuGlyAsnValLeuLeuLeuArgSerValSerThrAlaValTyrLys 149
Db 73 CGCTACCGCACTGGGCGACGCTCTCATCTCGCGCAGCTCAGCAGCGAGTCTACAG 132
Oy 150 ArgPheProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLys 169
Db 133 CGCTTCCCGACGCGCGACAGCTGTGTGAGTGAATCCGGGAGCCCTTACCTGCTCCAGAGC 192
Oy 170 GlnLeuGlnLysLeuSerLeuProHisAsnMetPheThrValProTyrValThrPheAla 189
Db 193 CAGTTGGAGAACTGAGCTTACACACACACATGTTCTGGGAGCCCTGGGTGGTGGTGGC 252
Oy 190 AsnLeuSerMetLysAlaTyrPheGlyGlyArgLysAspProIleLeuLeuGlnSer 209
Db 253 AACCTGTCATGTAAGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 312
Oy 210 LeuLeuAsnGlnMetAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrp 229
Db 313 CTGCTGACGAGATGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
Oy 230 IleSerIleProLeuValTyrThrGlnValAlaThrValAlaValTyrSerPhePheLeu 249
Db 373 ATTAGTGTCCCACTGCTGTATACACAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
Oy 250 ThrCysLeuValGlyArgGlnPheLeuAsnProAlaLysAlaTyrProGlnHisGlnLeu 269
Db 433 ACTTGTCTAGTGGGGGCGAGTTTCTGAACCCAGCCAGCCAGCCAGCCAGCCAGCCAG 492
Oy 270 AsnLeuValValProValPheThrPheLeuGlnPhePhePheTyrValGlyTyrPheLys 289
Db 493 GACCTCGTGTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
Oy 290 ValAlaGlnGlnLeuIleAsnProPheGlyLysAspAspAspPheGlnThrAsnTrp 309
Db 553 GTGGAGAGCAGCCATCATACCCCTTGGAGAGATGATGATGATGATGATGATGATGATG 612
Oy 310 IleValAspArgAsnLeuGlnValSerLeuLeuAlaValAspGlnMetHisGlnAspLeu 329
Db 613 ATTGTGACAGCAATTTTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
Oy 330 ProArgMetGluProAspMetTyrTyrAsnLysProGluLysProGluLysProTyrThala 349
Db 673 CCTCGAGATGAGCGGACATGCTGATGAATAGCCGAGAGAGAGAGAGAGAGAGAGAG 732
Oy 350 AlaSerAlaGlnPheArgAlaSerPheMetLysSerThrPheAsnIleSerLeuAsn 369
Db 733 GCTTCGCGCCAGATCCCTGAGCCTCTTATGGGCTCAACCTTCAACATAGAGCTGAGC 792
Oy 370 LysGlnLysMetGluPhe-----GlnProAsnGlnLysProGlnLysProAlaHisAla 387

```

```

Db 793 AACAGAGATGGAATTCACAGCCANTCCGGGAGAGAGAGAGAGATTCCTCCGCTG 852
Oy 388 GYLIIILGLYtrghe---LeuGlyLeuGln-SerHis-AspHisHis-PropoArg 405
Db 853 GCATCATTTGGCGGGTCTCCCAAAAGCCCTGACCCCCCATGATGATCATCTCCAGG 912
Oy 406 AlaAsnSer-----ArgThrLysLeuLeuTrp---ProLysArg 417
Db 913 GCCAAACATTCAGAGAACCCACCTTACTGTGGGCCCCCAAGA 957

RESULT 4
BE410951
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 773)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at: image.lml.gov
plate: LICM38 row: 1 column: 16
High quality sequence stop: 662.
FEATURES
source
1. 773
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3638175"
/clone_lib="NIH-MGC-21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(c). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT

204 a 239 c 176 g 154 t

Alignment Scores:

Pred. No.: 7.07e-106 Length: 773
Score: 1020.00 Matches: 195
Percent Similarity: 96.538 Conserved: 0
Best Local Similarity: 96.538 Mismatches: 6
Query Match: 32.698 Indels: 1
DB: 10 Gaps: 0

US-09-622-964-3 (1-585) x BE410951 (1-773)

```

Oy 384 AsnAlaHisAlaGlyIleIleGlyArgPheLeuGlnLeuGlnSerHisAspHisHis 403
Db 2 GATGCTACCTTACCATCATGATGGCGCTCTAGAGCTGAGTGCATGATCACCATCTCT 61
Oy 404 ProArgAlaAsnSerArgThrLysLeuLeuTrpProLysArgGlnSerLeuLeuGln 423
Db 62 CCCAGGGCAACATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 121

```

OY 424 GlyLeu-ProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnLysAspAsnLys 443
 DB 122 GGCCTGGCCCAAAAAACACAGAGGAGCCAAACAGACTTGGGGCCAGGAAGACACAA 181
 OY 443 SAATTPryLeuLysAlaValAlaPheLysSerGlyProLeuTyrGlnArgProG 463
 DB 182 GGCCTGGAGGCTTAAGGCTGTGGAGCCCTTCAGTCTCCCACTGATCAGAGGCCAGG 241
 OY 463 YTTTyrSerAlaProGlnThrProLysSerProThrProMetPhePheProLeuGluPr 483
 DB 242 CTACTACAGTCCCCACAGAGCCCTTCAGCCCTCCCTCCAGTCTTCCCTCCCTAGAAC 301
 OY 483 OSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLysSerLeu 503
 DB 302 ATCAGCCGCCCTCAAGCTTCACAGTGTCCAGGATGACACCAAGCAAAAGACTTAA 361
 OY 503 sThrValSerSerLysAlaLysLysSerPheGluLeuLeuSerGluSerAspGlyAlaLe 523
 DB 362 GACTGTGAGTCTGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCCGATGGGCCCTT 421
 OY 523 UmetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsnLeuThrAs 543
 DB 422 GATGAGACACCCCAAGATATCTCAAGTAGAGAGGAAAACTGTGAGTTTAACTGACGGA 481
 OY 543 PmePrGluIleProGluAsnHisLysGluProLeuGluGlnSerProThrAsnI 563
 DB 482 TATGCCAGAGATCCCAAAATCACCCTCAAGAAACCTTGGAAACATACCAACCAACAT 541
 OY 563 eHisThrThrLeuLysAspHisMetAspProTyrThrAlaLeuGluAsnArgAspGluAl 583
 DB 542 ACACACATCACTCAAAATCATCATGATCTTATTGGCCCTTGAAAAACAGGTCTGTCT 601
 OY 583 aHis 584
 DB 602 CCAC 605
 RESULT 5
 BE385296 593 bp mRNA linear EST 21-JUL-2000
 LOCUS 601277572F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3618665 5'
 DEFINITION mRNA sequence.
 ACCESSION BE385296
 VERSION BE385296.1 GI:9330661
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 593)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-femail.nih.gov
 Tissue Procurement: ATCC/DCMB/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: Image.Lnl.gov
 Plate: LNCM287 row: o column: 18
 High quality sequence stop: 593.
 FEATURES
 source
 1..593
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3618665"
 /clone_lib="NIH_MGC_20"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOT7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 176 a 177 c 132 g 108 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.9e-105 Length: 593
 Score: 1011.00 Matches: 192
 Percent Similarity: 97.46% Conservative: 0
 Best Local Similarity: 97.46% Mismatches: 4
 Query Match: 32.40% Indels: 1
 DB: 10 Gaps: 0
 US-09-622-964-3 (1-585) x BE385296 (1-593)
 OY 389 IleIleGlyArgPheLeuGlyLeuGlnSerHisAspHisProProAlaAsnSer 408
 DB 2 ATCATTTGGCCGCTTCTAGGCGCTGCAATCCATGATACCATCTCCACAGGCAAACTCA 61
 OY 409 ArgThrLysLeuLeuTyrProLysArgGluSerLeuLeuIleGlyLeu-ProLysAs 428
 DB 62 AGGACCAACTACTGTGGCCCAAGAGGAAATCCCTTCCACGAGGGCTGCCAAAAA 121
 OY 428 nHisLysAlaAlaLysGlnAsnValArgGlyGlnLysAspAsnLysAlaLysLeu 448
 DB 122 CCACAGGCGCCCAAGAGACCTTGAAGGCGCCAGAGAACAAAGCCCTGGAAGCTTAA 181
 OY 448 sAlaValAspAlaPheLysSerGlyProLeuTyrGlnArgProGlyTyrTyrSerAlaPr 468
 DB 182 GCGTGTGAGCGCTTCAAGTCTGCCACATCATACAGGCGACGACTACTAGTGGCCC 241
 OY 468 OGlnThrProLeuSerProThrProMetPhePheProLeuGluProSerAlaProSer 488
 DB 242 ACAGAGCCGCCCTCAGCCCTCCATGTTCTTCCCTGAAACATACAGCGCGTCAA 301
 OY 488 sLeuHisSerValThrGlyIleAspThrLysAspLysSerLeuLysThrValSerSerG 508
 DB 302 GCTTTCACAGTGTCAACAGGCTATACACCAAGAACAAAGCTTAAAGCTGTGAGTCTGG 361
 OY 508 yAlaLysLysSerPheGluLeuLeuSerGluSerAspGlyAlaLeuUmetGluHisProG 528
 DB 362 GGCACAAAGAAAGTTTGAATGTCTTCAGAGACGATGGGCCCTTATGAGACACCCAGA 421
 OY 528 vAlaSerGlnValArgArgLysThrValGluPheAsnLeuThrAspMetProGluIlePr 548
 DB 422 AGTATCTCAAGTAGAGAGGAAACCTGTGAGTTTAACTGACGATATGCCAGAGATCCC 481
 OY 548 OGluAsnHisLysLeuGluProLeuGluGlnSerProThrAsnIleHisThrThrLeu 568
 DB 482 CGAAATTCACCTCAAGAAACCTTGGAAACATACCAACCAACATACACTACACTCA 541
 OY 568 sAspHisMetAspProTyrThrAlaLeuGluAsnArgAspGluAlaHis 584
 DB 542 AGATCACAGATGATCTTATTGGCCCTTGGAAAAACAGGTCTGTCTCCAC 590
 RESULT 6
 B1756228 735 bp mRNA linear EST 25-SEP-2001
 LOCUS 603024265F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194649 5'
 DEFINITION mRNA sequence.
 ACCESSION B1756228
 VERSION B1756228.1 GI:15747806
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11487 row: a column: 18
High quality sequence stop: 577.
Location/Qualifiers
1. 735

FEATURES
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:5194649"
/clone_lib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI,
Site: 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."
BASE COUNT 135 a 229 c 214 g 157 t
ORIGIN

Alignment Scores:

Pred. No.: 5,35e-98 Length: 735
Score: 951.00 Matches: 198
Percent Similarity: 90.27% Conservative: 6
Best Local Similarity: 87.61% Mismatches: 8
Query Match: 30.48% Indels: 14
DB: 13 Gaps: 2

US-09-622-964-3 (1-585) x B1756228 (1-735)

QY 52 leuAlaLeuThrGluGlnGlnGlnLeuMetPheGlnLysLeuThrLeuTyrCysAspSer 71
DB 60 CTGGCCCTCAGGAAACAGACGCTGATGAGAACTGACTCTGATTGGAGAC 119
QY 72 TyrIleGlnLeuLeuPheProIleSerPheValLeuGlyPheTyrValThrLeuValValThr 91
DB 120 TACATCCAGCTCATCCCATTTCTTCGTCGCTGCGCTTACGTCGAGCGTCTGAC 179
QY 92 ArgTTrpAsnGlnTyrGlnAsnLeuProTTrpProAspArgLeuMetSerLeuValSer 111
DB 180 CGCTGTGGAAACAGTACGAGACCTGCGTGGCCGACCGCTCATGAGCTGGTGTG 239
QY 112 GlyPheValGlnGlyLysAspGlnGlnSerArgLeuLeuArgTThrLeuLeuArgTyr 131
DB 240 GCGTTGTGAAAGGCAAGAGACGAGCCGCTCTGCGGCGACGCTCATCCCTAC 299
QY 132 AlaAsnLeuGlyAsnValLeuLeuLeuArgSerValSerThrAlaValTyrLysArgPhe 151
DB 300 GCCAACCTGGGCAAGCTGCTCATCCGCGACGCTACAGACCGCACTTACAGCCCTTC 359
QY 152 ProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeu 171
DB 360 CCCACGCGCCACACCTGCTGCAAGCAGGCTTATGACTCGCGGAGAACACAAACAGCTTG 419
QY 172 GlnLysLeuSerLeuProHisAsnMetPheTTrpValProTTrpValTTrpPheAlaAsnLeu 191
DB 420 GAGAAACAGGAGCTTACACACAAATGTTCTGGGTGCTGCTGGTGTGCTTCCAACTGG 479
QY 192 SerMetLysAlaTTrpLeuGlyGlyArgLysArgPheProIleLeuLeuInsSerLeuLeu 211
DB 480 TCAATGAAGGGGTGCTGAGAGTCGAATCCGGGAGACCCATCTCTGCTCCAGAGCCCTGCTG 539

FEATURES
Source

QY 212 AsnGlnMetAsnThrLeu-ArgThrGlnCysGlyHisLeuTyrAla-TyrAspTrpIles 231
DB 540 AACGAGATACACACCTGGCGCTTCGCTGCTGAGACCTGTATGCTTAACGATGATTA 599
QY 231 erIleProLeuValTyrThr-GlnValVal-----ThrValAlaValTyr 245
DB 600 GTATCCCTC-----ACTGGGTGTATCCCGGGGTGGTGGTGGTGGTGGTGGTGG 650
QY 246 SerPhePheLeuThrCys-LeuValGlyArg-GlnPheLeuAsn-ProAlaLysAlaTyr 264
DB 651 GGTCTTCTTCCTGAGCTGTCTTGTGAGGCGGCGGCTTGTGAACCCGCGACGCTTC 710
QY 265 ProGly 266
DB 711 CCTGGG 716

RESULT 7
LOCUS BM691456
DEFINITION UI-E-C11-abe-f-10-0-UI-r1 UI-E-C11 Homo sapiens cDNA clone
ACCESSION BM691456
VERSION BM691456.1 GI:19004714
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 537)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. 537
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="UI-E-C11-abe-f-10-0-UI"
/clone_lib="UI-E-C11"
/issue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pTrp3-Pac (pharmacia) with a
modified polylinker; Site: 1: EcoR I; Site: 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTrp3-Pac Vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

BASE COUNT 133 a 172 c 139 g 93 t
 ALIGNMENT Scores:
 Pred. No.: 4.44e-97 Length: 537
 Score: 941.00 Matches: 174
 Percent Similarity: 98.86% Conservative: 0
 Best Local Similarity: 98.86% Mismatches: 2
 Query Match: 30.16% Indels: 0
 DB: 14 Gaps: 0

US-09-622-964-3 (1-585) x BM691456 (1-537)

OY 306 GluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeuAlaValAspGluMet 325
 |||||
 DB 8 GAGACCAACTGATGTCGACAGAAATTGCAAGTGTCCCTGGCTGGCTGATGACATG 67
 |||||

OY 326 HisGlnAspLeuProAlaGlyMetGluProAspMetTrpAsnLysProGluProGlnPro 345
 |||||
 DB 68 CACCAGACCTGCTCGATGAGACCGGACATGTACTGATTAAGCCGACACACAGCC 127
 |||||

OY 346 ProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMetGlySerThrPheAsn 365
 |||||
 DB 128 CCGTACACAGCTGCTTCCGCCAGTTCCTGCAAGCTCTTATGGCTCCACCTCAAC 187
 |||||

OY 366 IleSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnAlaAspGluAlaAspAla 385
 |||||
 DB 188 ATCGCCTGAACAAGAGAGATGAGTGCAGCCCATCAAGAGAGAGAGAGATGCT 247
 |||||

OY 386 HisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAspHisSerProArg 405
 |||||
 DB 248 CAGCTGATCATATGGCCCTTCCATGAGCTTCCATGATGATCAATCTCCAGG 307
 |||||

OY 406 AlaAsnSerArgThrLysLeuLeuTrpProLysArgLysLeuLeuHisGluGlyLeu 425
 |||||
 DB 308 GCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGAATCCCTTCCAGAGGCTG 367
 |||||

OY 426 ProLysAsnHisLysAlaAlaLysGlnAsnValArgLysGlnLysAsnLysAlaTrp 445
 |||||
 DB 368 CCAAAAACCAAGAGAGAGCAAGAACGTTAGGGCCAGAGAACCAAGAGGCTG 427
 |||||

OY 446 LysLeuLysAlaValAspAlaPheLysSerGlyProLeuTrpGlnArgProGlyTyr 465
 |||||
 DB 428 AACCTTAAGCTGTGACGCTTCAAGTCTGCCCATCTATCAAGAGGCTACTAC 487
 |||||

OY 466 SerAlaProGlnThrProLeuSerProThrPheMetPhePheProLeu 481
 |||||
 DB 488 AGTGGCCCAAGAGAGGCTCCAGCCCACTGCATGTCTTCCCTCA 535
 |||||

RESULT 8
 BQ436824 930 bp mRNA linear EST 24-MAY-2002
 LOCUS AGENCOUPT_7583107 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6049729
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ436824
 VERSION BQ436824.1 GI:21175900
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaas-r@mail.nih.gov
 Tissue Procurement: ATCC/DCPD/DPF
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILINT)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILINT at:

http://image.llnl.gov
 Plate: LRAM1300 row: n column: 02
 High quality sequence stop: 450.
 Location/Qualifiers
 1..930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6049729"
 /clone_id="NIH_MGC_72"
 /issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1; Not:
 /note:2: Salt: Cloned unidirectionally. primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 237 a 309 c 217 g 166 t 1 others
 ORIGIN

ALIGNMENT Scores:
 Pred. No.: 5.42e-96 Length: 930
 Score: 935.00 Matches: 202
 Percent Similarity: 78.34% Conservative: 15
 Best Local Similarity: 72.92% Mismatches: 37
 Query Match: 29.97% Indels: 23
 DB: 14 Gaps: 7

US-09-622-964-3 (1-585) x BQ436824 (1-930)

OY 295 IleAsnProPheGlyGluAspAspAspPheGluThrAsnTrpIleValAspArgAsn 314
 |||||
 DB 100 ATCAACCCCTTTGAGAGAGATGATGATTTGAGACCAACTGATGTCGACAGAAAT 159
 |||||

OY 315 LeuGlnValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProAlaGlyMetGluPro 334
 |||||
 DB 160 TTCAGAGTGTCTGTTGGCTGTGATGAGATGCACAGAGCTGCTCGATGAGAGCCG 219
 |||||

OY 335 AspMetTrpTrpAsnLysProGluProGlnProTyrThrAlaAlaSerAlaGlnPhe 354
 |||||
 DB 220 GACATGTACTGGAATAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 279
 |||||

OY 355 ArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGluGluMetGlu 374
 |||||
 DB 280 CGTCGAGCTCTCTTGTGGCTCCACCTCAACATCAGCTGMAACAAGAGAGAGAGAG 339
 |||||

OY 375 PheGlnProAsnGlnLysAspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeu 394
 |||||
 DB 340 TTCACGCCCAATCAG 399
 |||||

OY 395 GlyLeuGlnSerHisAspHisAspProArgAlaAsnSerArgThrLysLeuLeuTrp 414
 |||||
 DB 400 GGCCTGACATGCCATGATCACCATCTCCAGAGGCAAACTCAAGAGCAAACTACTGTG 459
 |||||

OY 415 ProLysArgGlnSerLeuLeuHisGluGlyLeuProLysAsnHisLysAlaLysGln 434
 |||||
 DB 460 CCCAAGAGGGAATCCCTTCCACAGAGGCTGCCCCAAAACCAAGAGAGAGAGAGAG 519
 |||||

OY 435 AsnValArgGlyGlnLysAsnLysAlaTrp-LysLeuLysAlaValAspAlaPheLys 454
 |||||
 DB 520 AACGTAGAGGCGCAGAGAACACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 579
 |||||

OY 454 SerGlyProLeuTrpGlnArgPro-GlyTyrTyrSer-AlaProGln-ThrProLeuSe 473
 |||||
 DB 580 GTCCTCCCACTGATCAAGAGCCATGCTTACTAGTGGCCCCCAGAGAGAGAGAGAG 639
 |||||

OY 473 rPro--ThrPrometPhePheProLeuGlnProSer--AlaProSerLys----- 488
 |||||

OY 640 CCCCATTCCCATATTTCTTCCCTTAAACATACGAGGCGCTCAAAAGCTTCCC 699
 |||||

OY 489 ----LeuHisSerValThrGlyIleAspThrLysAspLysSerLeuLysThrValSer-- 506
 |||||

OY 700 ACGCGTGCACGAGGCATAGAAACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 759
 |||||

OY 507 -----SerGlyAlaLysLysSerPhe-GluLeuLeu-----SerGluSerAspGlyA 522
 |||||

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	227	CTCTGGTACTACATCATCCGCTTATTATTAAAGCTG6CCCTCACGGAGAACACAGCTG					
Qy	61	MetPheGluLysLeuThrLeuIleuTyrCysAspSerTyrIleGlnIleuIleProIleSerPhe					
Db	287	ATGTTTGAGAAATGACTGCTGTATTCGACAGACTACATCATCGAGCTCATCCCATTTCTTC					
Qy	81	ValLeuGlyPheTyrValThrLeuValIleThrArgTyrTrpAsnGlnTyrGluAsnLeu					
Db	347	GTCCTGGGCTTCTACGTACGCGCTGGTGGTGGACCCGCTGGGAAACGATACGAACTCG					
Qy	101	ProTyrProAspArgLeuMetSerLeuValIleSerGlyPheValGluIleLysAspGluGln					
Db	407	CCGTGGCCCGACCGGCTCATGTAGACCTCGTGGTGGGCTTCTGTGGAGCGAAGCAAGCA					
Qy	121	SerArgLeuLeuLeuArgArgThrLeuLeuIleArgTyrIleAsnLeuGlyAsnValLeuIle					
Db	467	GGCCGGGCTCTCTGGGGGACGCTCATCCGCTTACCGGCAACCTTGGGCAACCTGCTCAT					
Qy	141	ArgSerValSerThrAlaValTyrTyrAspArgPheProSerAlaGlnHisLeuValGlnIle					
Db	527	CCGACGGTACAGACCGGACGCTTACACAGCGCTTCCACG - GCCCAGACACTGGTGCAGCA					
Qy	161	GlyPheMetThrProAlaGlnHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet					
Db	586	GGCTTATATGACTACGCGGACAGAACACAGAGCTGGAGAAACTGAGCTACACACAACTG					
Qy	181	PheTrpValProTyrP					
Db	646	TTCGTGGTGGCCCTGG 660					
RESULT 10							
LOCUS	BM685501	526 bp	mRNA	linear	EST 27-FEB-2002		
DEFINITION	UI-E-C10-aad-h-10-0-UI-r1 UI-E-C10 Homo sapiens cDNA clone						
ACCESSION	BM685501						
VERSION	BM685501.1	GI:18995397					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
REFERENCE	1 (bases 1 to 526)						
AUTHORS	Bonaldo,M.F., Lennom,G. and Soares,M.B.						
TITLE	Normalization and subtraction: two approaches to facilitate gene						
	discovery						
	Genome Res. 6 (9), 791-806 (1996)						
JOURNAL	9704447						
MEDLINE							
COMMENT	Contact: Soares, MB						
	Program for Rat Gene Discovery and Mapping						
	University of Iowa						
	451 Eckstein Medical Research Building Iowa City, IA 52242, USA						
	Tel: 319 335 8250						
	Fax: 319 335 9565						
	Email: msoares@blue.weeg.uiowa.edu						
	Tissue Procurement: Dr. Gregg Hageman						
	cDNA library preparation: Dr. M. Bento Soares, University of Iowa						
	cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa						
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa						
	Clone Distribution: Researchers may obtain clones from Research						
	Genetics (www.resgen.com).						
	Seq primer: M13 Reverse.						
FEATURES	location/Qualifiers						
source	1..526						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
	/clone="UI-E-C10-aad-h-10-0-UI"						
	/clone_11b="UI-E-C10"						
	/tissue.type="RPE and Choroid"						
	/dev.stage="adult"						
	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"						
	/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a						
	modified polylinker; Site_1: EcoR I; Site_2: Not I;						

OY		237	TGtGGnAlValTThVAlAAlValTYrSerPhePheLeuThrCysLeuValGIyArgIn	256
Db		493	ACCCAGCTAGTCACACTACACTGTACACCTACTCTTGCTCATCGGCCCTCAg	552
OY		257	PheLeuAsnProAlAlLysAlATyrProGlyHisGLuLeuAspLeuValVal-ProValPh	276
Db		553	TTCCTAGACCCTGCACAGGGCTACAAGAACCACACCCTGACCTATGGCTACCCCATT	612
OY		276	eThrPheLeuGlnPhePhePheTyrValGLyTrpLeuTySValAlAlGlnLeuIleAs	296
Db		613	CACCTGCTGCAGTCTCTCTTACGCTGGCTGCTTAAGTAGACAGACACACTATATPA	672
OY		296	nProPheGLyluAspaSPasPaSPheGluThrAsnTrpIleValAspArgAsnLeucI	316
Db		673	-CCCTTTGGAGAGGACGAAGACGACTTTGAGACCAATT-CTTATTGACCCAAATAATTC	730
OY		316	nValSerLeuLeuAlaValASPglUmeHisGlnAspLeuProArgMetGluProaspNe	336
Db		731	GGTGTCAATGCAGCTGTAAAGTATGAATATTTATAGACAGCTGGGCCAGGCAAAAAGATCT	790
OY		336	t 336	
Db		791	A 791	

RESULT 12	
AA307119	
LOCUS	AA307119
DEFINITION	EST118031 Colon carcinoma (HCC) cell line Homo sapiens CDNA 5' end,
ACCESSION	AA307119
	469 bp mRNA linear EST 18-APR-1997
	mRNA sequence.

ACCESSION AA307119 GI:1959449
VERSION AA307119.1
KEYWORDS EST.

REFERENCE
ADITHYAN
ADITHYAN, A. R., and E. J. S. 1964. The
1 (bases 1 to 469)
Adams, M. D., Kerlavage, A. R., and E. J. S. 1964. The
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 469)
Adams, M. D., Kerlavage, A. R., and E. J. S. 1964. The
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 469)

TITLE	Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL	Nature 377 (6547 Suppl.), 3-174 (1995)
COMMENT	96026280
OTHER ESTS	THC123650

FEATURES	Location/Qualifiers
source	1. .469
	/organism="Homo sapiens"

```

/db_xref="ATCC:(host):108624"
/db_xref="taxon:9606"
/cclone_1hp="Colon carcinoma (HCC) cell line"
/cusue_type="colon"
/cell_type="KM12C"
/cell_line="KM12C(HCC)-parental human colon carcinoma
;Dukes B2"
/note="Organ: colon; Vector: pBluescript SK-; Site:1;"
ECCORI: Site_2: XhoI"
BASE COUNT      131 a      141 c      116 g      80 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:
Score:      1,71e-82      length:      469
Percent Similarity:      813.00      Matches:      154
Best Local Similarity:      98.72%      Conservative:
Query Match:      98.72%      Mismatches:      2
DB:      26.06%      Indels:      0
      9      Gaps:      0

```

US-09-622-964-3 (1-585) x AA307119 (1-469)

QY 378 AsnGlnGlnAspIuGluAspAlaHisAlaGlyIleIleGlyArgPheIuGluGln 397
Db 2 AATCAGAGAGACGAGAGATGCTCAACGCTGGCATCTTGGCGCTTCTAGCGCTTCAG 61
QY 398 SerHisAspHisHisProProAlaAlaAsnSerArgThrLysLeuLeuTrpProlLysArg 417
Db 62 TCCCAAGATCACCATCTCTCCACAGGCCAAACTCAAGGACCAACTACTGTGGCCCAAGAG 121
QY 418 GluSerLeuLeuHisGluGlyLeuProlLysAsnHisLysAlaAlaLysGlnAsnValArg 437
Db 122 GAATCCCTTCTCCACAGAGGCTGCCCCAAACCAACAAGCACCACCAACGAAGCTTAGG 181
QY 438 GlyGlnGlnAspAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLysSerGlyPro 457
Db 182 GCCCAGGAAGACACAACAGGCTGGAAGCTTAAAGCGTGTGAGCCCTTCAAGTGTGCCCA 241
QY 458 LeuTrpGlnArgProGlyTrpTrpLysSerAlaProGlnThrProLeuSerProThrPromet 477
Db 242 CTGATCTCAAGAGCCAGGCTACTACAGTGGCCCAAGACACCCCTCAGCCCATCCCATG 301
QY 478 PhePheProLeuGluProSerAlaProSerLysLysHisSerValThrGlyIleAspThr 497
Db 302 TTCTTCCCCCTAGAACCAATCAGGCGCGTCAAAAGCTTCACAGTGTCAACAGGCATACAGACC 361
QY 498 LysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSer 517
Db 362 AAAGCAAAACCTTAAACACTGTGAGTCTGTGGGCCCAAGAAAGTTTGANTTGGCTCTCA 421
QY 518 GluSerAspGlyAlaLeuMetGlnHisProGluValLysGlnValArg 533
Db 422 GAGAGCAATGGGCGCTTATGTAGAGCACCCACAGAAAGTATGTCAAGTAGG 469

LOCUS	455 bp	mRNA	linear	EST 24-JAN-2001
LOCUS	BG013943			
DEFINITION	IL5-GM0240-201200-347-b04	GM0240	Homio sapiens	CDNA, mRNA sequence.
ACCESSION	BG013943			
VERSION	BG013943.1	GI:12464642		
KEYWORDS	EST.			
SOURCE	human.			

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags
JOURNAL
MEDLINE
20202663
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ptl=IL56t2-IL5-GN0240-
201200-347-604&t3=2000-12-20&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 390.

FEATURES

source
1.455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0240"
/dev_stage="Adult"
/note="Organ: placenta, normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
Application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 98 a 139 c 120 g 97 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 7.91e-82 Length: 455
Score: 807.00 Matches: 149
Percent Similarity: 98.688 Conservative: 0
Best Local Similarity: 98.688 Mismatches: 2
Query Match: 25.878 Indels: 0
DB: 12 Gaps: 0
US-09-622-964-3 (1-585) x BG013943 (1-455)

260 PROALALYALATYRPROGLYHISGLULALAPLEUVALPROVALPHERPHELEU 279
2 CCAAGCCAGGCTACCTGGCCATGAGCTGAGCTGCTGCTCCCTTTCACGTTCTG 61
280 GLNPHEPHEPHERYVALGTYRPLEULYSVALALAGLGLNLEULLEASNPROPHESLY 299
62 CAGTCTCTCTCTATGTTGGCTGGCTGAGAGTGCGACACCTCATCAACCCCTTTGCA 121
300 GLUASPASPASPAPHEGLUTHRASNTRPILLEVALASPARGASNULEGLNVALSERLEU 319
122 GAGGATGATGAGATTTTGAGACCACTGATTCGACAGAAATTTGACAGGTGCCCG 181
320 LEUALAVALASPLMEHISGLINASPLEUPROARGMETGLUPROASPMETTYRTPASN 339
182 TTGGCTGTGATGATGATGACACAGGACCTGCTCGATGAGACCGGACATGTACTGGAT 241
340 LYSPTROGLUPROGLNPROPTOTYRTHRALAASERALAGLNPHARGARGALASERPHE 359
242 AAGCCCGAGCCACACCCCTTACACAGCTGCTCCGCCAATTCGTCGACCTCTTT 301
360 METGLYSERTHRPHASNILESERLEUASNLYSGLUGLUMETGLUPHEGLNPROASGLN 379
302 ATGGGCTCCACCTTAAACATCAGCTGAACAAGAGGAGATGAGATTCACGCCCAATCAG 361
380 GLUASPGIUGLUASPALAHISALAGLYILEIAGLYARGPHELEUGLYLEUNGINSERTHS 399
362 GAGGACGAGGAGGAGATCTCACGCCCTGCATCATNGCCGCTCTCTAGGCTGAGTCCAT 421
400 ASPHISHSPPTROARGALASNSERARGTHR 410
|||||

Db 422 GATCACCATCTCTCCAGGCGCAACTCAGACGCC 454

RESULT 14
BM707649 522 bp mRNA linear EST 28-FEB-2002
LOCUS
DEFINITION
UI-E-C11-afg-o-15-0-UI-r1 UI-E-C11 Homo sapiens cDNA clone
UI-E-C11-afg-o-15-0-UI 5', mRNA sequence.
BM707649
VERSION
BM707649.1 GI:19020907
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 Reverse

FEATURES

source
1.522
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UI-E-C11-afg-o-15-0-UI"
/clone_lib="UI-E-C11"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(d718 tail). The sequence tag for this library is ACCCT.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."
BASE COUNT 100 a 182 c 130 g 110 t
ORIGIN

Alignment Scores:
Pred. No.: 4.02e-80 Length: 522
Score: 793.00 Matches: 153
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 1
Query Match: 25.428 Indels: 0
DB: 14 Gaps: 0

US-09-622-964-3 (1-585) x BM707649 (1-522)

QY 1 Methrliethrtyrthrsrclnvalaalsnlaarglenglyserpheserrargleu 20
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 14:57:56 ; Search time 347.053 Seconds
(without alignments)
3477.440 Million cell updates/sec

Title: US-09-622-964-3

Perfect score: 3120
Sequence: 1 MTTTYSQVANNRLGSRSL.....TLKDHPYALNREDAHS 585

Scoring table:

BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1439767 segs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-rlh
-Q/cg2_1/USPTO.spool/US0622964/runt_22072003_101120_26142/app_query.fasta_1.2069
-DB-published.Applications.NA -QEXT-fastap -SUFFIX-rmpb -MINMATCH-0.1
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-biosum62
-TRANS-human40.cdi -List-45 -LOCAL-ALIGN-200 -THR SCORE-pct -THR MAX-100
-THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pico -NORM-ext -HEARSIZE-500 -MINLEN-0
-MAXLEN-200000000 -USER-US09622964 @CGN_1_1_131 -runt_22072003_101120_26142
-NCPU-6 -ICPU-3 -NO MAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSBLOCK-100
-LONGLOG -DEV TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAEXT-0.5 -DELOP-6 -DELEXT-7

Database :

1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/2/pubpna/PT_NEW_PUB.seq:*
3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/2/pubpna/PTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
11: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
12: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
15: /cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result Query 8
No. Score Match Length DB ID Description
1 1571 50.4 1263 12 US-09-746-783-3 Sequence 3, Appl1

SUMMARIES

2	785.5	25.2	1292	15	US-10-198-846-11070	Sequence 11070, A	
C	662	21.2	1350	10	US-09-768-826-16	Sequence 16, Appl	
C	4	333.5	10.7	748	15	US-10-027-632-134529	Sequence 134529, A
5	304	9.7	853	15	US-10-198-846-11346	Sequence 11346, A	
C	6	300	9.6	578	15	US-10-027-632-134530	Sequence 134530, A
C	7	264	8.5	462	15	US-10-198-846-7861	Sequence 7861, Ap
C	8	264	8.5	615	15	US-10-198-846-7843	Sequence 7843, Ap
9	238	7.6	507	12	US-09-764-872-195	Sequence 195, App	
10	171	5.5	751	15	US-10-198-846-1862	Sequence 1862, App	
C	118	3.8	3186778	15	US-10-027-632-17961	Sequence 17961, A	
12	114.5	3.7	1630	11	US-09-925-300-141	Sequence 141, App	
13	111.5	3.6	3280	11	US-10-171-581-275	Sequence 275, App	
C	14	111.5	3.6	1276	15	US-10-114-893-87	Sequence 87, Appl
C	15	110.5	3.5	609	15	US-10-027-632-245664	Sequence 245664, A
16	108	3.5	4152	15	US-10-226-315-5	Sequence 5, Appl1	
17	108	3.5	5294	15	US-10-226-315-1	Sequence 1, Appl1	
18	105.5	3.4	3067	10	US-09-925-301-580	Sequence 580, App	
19	105	3.4	2145	15	US-10-037-270-421	Sequence 421, App	
20	104.5	3.3	2361	10	US-09-815-242-6006	Sequence 6006, Ap	
C	21	104	3.3	654	11	US-09-738-626-2901	Sequence 2901, Ap
22	104	3.3	3309400	11	US-09-738-626-1	Sequence 1, Appl	
23	103.5	3.3	1362	15	US-10-156-761-4828	Sequence 4828, Ap	
24	103.5	3.3	9025608	15	US-10-156-761-1	Sequence 1, Appl	
25	102	3.3	1973	11	US-09-903-396A-1	Sequence 1, Appl	
26	101.5	3.3	1902	11	US-09-801-368-337	Sequence 337, App	
27	100.5	3.2	18404	15	US-10-142-515-10	Sequence 10, Appl	
28	100	3.2	2430	15	US-10-156-761-2531	Sequence 2531, Ap	
29	100	3.2	4143	15	US-10-226-315-6	Sequence 6, Appl1	
30	100	3.2	5350	15	US-10-226-315-3	Sequence 3, Appl1	
31	99.5	3.2	1600	12	US-09-866-050A-553	Sequence 553, App	
32	99.5	3.2	1600	15	US-10-152-661-553	Sequence 553, App	
C	33	3.2	9025608	15	US-10-156-761-1	Sequence 1, Appl	
34	99	3.2	2142	11	US-09-801-368-407	Sequence 407, App	
C	35	98.5	3.2	2121	11	US-09-940-766-1	Sequence 1, Appl1
C	36	98.5	3.2	2121	15	US-10-142-174-32	Sequence 32, Appl
C	37	98.5	3.2	2121	15	US-10-142-174-36	Sequence 36, Appl
C	38	98.5	3.2	2140	15	US-10-142-174-14	Sequence 14, Appl
39	98.5	3.2	2576	11	US-09-880-107-2138	Sequence 2138, Ap	
40	98.5	3.2	2576	11	US-09-942-429A-6	Sequence 6, Appl1	
C	41	98.5	3.2	12047	15	US-09-764-847-1297	Sequence 1297, Ap
C	42	98.5	3.2	12047	15	US-10-092-154-1397	Sequence 1397, Ap
C	43	98.5	3.2	15987	11	US-09-764-847-1301	Sequence 1301, Ap
C	44	98.5	3.2	15987	15	US-10-092-154-1301	Sequence 1301, Ap
45	98	3.1	1324	15	US-10-149-819-26	Sequence 26, Appl	

ALIGNMENTS

RESULT 1
US-09-746-783-3
Sequence 3, Application US/09746783
Publication No. US2003004935A1
GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John M.

LaValle, Edward R.

Raele, Lisa A.

Treacy, Maurice

Spaulding, Vikki

Agostino, Michael J.

Howes, Steven H.

Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/746,783
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MilesInc, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Alignment Scores:
Pred. No.: 7,32e-171 Length: 1263
Score: 1571.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 1
Query Match: 50.35% Indels: 0
Gaps: 0

US-09-622-964-3 (1-585) x US-09-746-783-3 (1-1263)
QY 292 GIUGLNULEUASPRPHGGLUASPRASPRASPRHGLUTHRASNTPLLEVAL 311
DB 1 GAGCAGCATCATACCCCTTTGAGAGGATGATGATTTGAGAACCAACGATTTGC 60
QY 312 ASPARGASNULEUINVALSERLEUENVALAVALASPLUMETHISGLNASPLEUPROATG 331
DB 61 GACAGAGATTGTCAGAGTGTCCCTGTGGCTGTGATGATGATCACCAGAGACTGCTCGG 120
QY 332 METGLUPROASPMETTYTTPRANLYSPROGLUPROGLNPROPTOTYTHRALAASER 351
DB 121 ATGGAGCCGACATGTACTGAGATTAAGCCCGAGCCACAGCCCCCTACACAGCTCTCC 180
QY 352 ALAGINPHARGARGALASERPHMETGLYSERTPHRAENLISERLEUASNUYGLU 371
DB 181 GCCCAGTCCGTCGAGCTCTTTATGGGCTCCACCTTCACATCAGCTGAAACAAAGAG 240
QY 372 GLUMETGLUPHEGLNPROASNGLNUASPLUGLNUASPRALANISALGLYLETLEGLY 391
DB 241 GAGATGGAGATTCAGACCCCAATCAGAGAGAGAGAGATGCTCAGCTGGCATCTTGGC 300
QY 392 ARGPHLEUUGLYLEUGLNLSERHISAPHLISHPROPTOALGALAASNSEATGTHLYS 411
DB 301 CGCTTCAGAGCTGCAATCCATCATCATCTCCCAAGGCAAACTCAAGAGACAA 360
QY 412 LEUENUTPROLYSARGGLUSERLEUENHISGLUGLYLEUPROLYSASNHISLYALA 431
DB 361 CTAGCTGGCCCAAGAGAGATCCCTTCCACAGAGGCTGCCCCAAAACACAGAGCA 420
QY 432 ALALYSGINASNVALARGGLYGLNLUASPRANLYSALATPLYSLYUVALAVALASPR 451
DB 421 GCCAAACAGAACGTTAGGGCCAGAGAACAAAGGCTTGAAAGCTTAAGCTGTGGAC 480
QY 452 ALAPHELYSSERGLYPROLEUTYGLNARGPROGLYTYTYTYSERIALAPROGLNTHPRO 471
DB 481 GCCTTCAGATCTGCCCACTGTATCAGAGGCCAGGCTACTAGTGGCCCAAGAGGCC 540
QY 472 LEUSERPROTHPROMETPHEPHEPROLEUGLUPROSERIALAPROSERLYLEUENHISER 491
DB 541 CTCAGCCCACTGCCATGTTCTTCCCTTAGAACCATCAGCCGCTGAAAGCTTCACAGT 600

QY 492 VALTHRGYLLIENASPHRILYSAPLYSSERLEULYSTHVALSERSEGLVALALYSLYS 511
DB 601 GTCCAGGCGATGACACCCAAAGCAAAAGCTTAAGACGTGTGAGTCTGGGCGCAAGAA 660
QY 512 SERPHEGLULEUENSERGLUSERASPLUGLYALALEUMEGLUHISPROGLUVALSERGLN 531
DB 661 AGTTTGAATTTGCTCTCCAGAGAGCGATGGGCTTGATGAGACCCCAAGATCTCA 720
QY 532 VALARGLYSTHVALGLUPHEASNLEUTHRASPMEPROGLULIETPROGLUASNHIS 551
DB 721 GTGAGAGAGAAAACGTGGAGTTAACCTGACGAGGATATCCAGAGATCCCGCAATTCAC 780
QY 552 LEULYSGLUPROLEUGLUSERPROTHASNLHISHTHTRLEULYSAPHLISMET 571
DB 781 CTCAAAGAACCTTTGGAACAAATCACCACCAACATACACACTCAAGATCACATG 840
QY 572 ASPROTYTTPRALALEUGLUNASPRASPLUGLYALANISER 585
DB 841 GATCCTTATTTGGGCTTGAGAAACAGGAGTGAAGCATCTCC 882

RESULT 2

US-10-198-846-11070
Sequence 11070, Application US/10198846
Publication No. US20030099974A1

GENERAL INFORMATION:

APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1287, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Alignment Scores:

Pred. No.: 3,11e-80 Length: 1292
Score: 785.50 Matches: 150
Percent Similarity: 62.99% Conservative: 27
Best Local Similarity: 53.38% Mismatches: 53
Query Match: 25.18% Indels: 51
Gaps: 1

US-09-622-964-3 (1-585) x US-10-198-846-11070 (1-1292)

QY 82 LEUGLYPHELYVALTHRLEUVALVALTHRARGTRTPRASNGLINTYRGUASNULEUPRO 101
DB 317 ATAGCGTTTATGTACTCTGTGTAACCAAGGAGTGGAGAACAGTTTGTAATTTGCC 376
QY 102 TTPROASPARGLUENMETSERLEUVALSERGLYPHEVALGLUGLYLSASPGLUGLINSER 121
DB 377 TGGCAGACAGGCTAATGTTCTCATCTCTAGCAGTGTTCAGGAGAGGAGACAGGG 436
QY 122 ARGLEUENARGATGTHRLEULIENARGTYVALASNULEUGLYASNVALLEULIENARG 141
DB 437 CGCTGCTTGAAGGAGCGTGAATGCGCTACGTCAATCTCACTCCGCTCATCTTCCG 496
QY 142 SERVALSERTHRALVALTYLYSARGPHEPROSERALAGINHISLEUVALGLNALGLY 161


```
Db 497 TCGGTGACGACTGCTGTGTACAAAAGATTCCACCAATGACGACCTGGTGTGAAGCAGGT 556
Qy 162 PhenethrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetPhe 181
Db 557 TTTATGACACAGATGAAGGAATTTATTCACACACCTCAGTCTCTCATCTGAATAT 616
Qy 182 TTPValProTrpValITrPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArgIle 201
Db 617 TGGGTTCATTCATCTGTGTGGAATCTTGCACTAAAGCCGGAATGAGATGATC 676
Qy 202 ArgAspProIleLeuGlnSerLeuLeuAsnGlnMetAsnTrpLeuArgTrpGlnCys 221
Db 677 AGAGACAGCTGTGTATCTGCATCTGTATGATGATGATGATGATGATGATGATGAT 736
Qy 222 GLYHisLeuTrpAlaTrpIleSerIleProLeuValITrPheGlnValValThr 241
Db 737 AGCCTCTTATTCGTTATGACTGGTGGATTCGCTGTTACACCCAG----- 787
Qy 242 ValAlaValITrSerPhePheLeuTrpCysLeuValGlyArgGlnPheLeuAsnProAla 261
Db 787 ----- 787
Qy 262 LysAlaTrpProGlyHisGluLeuAspLeuValProValPheTrpPheLeuGlnPhe 281
Db 787 ----- 787
Qy 282 PhePheTrpValGlyTrpLeuLysValAlaGluGlnLeuIleAsnProPheGlyGluAsp 301
Db 788 -----GTAGCAGAGCAGACTTATCAACCTTTTGGAGAAAGAT 823
Qy 302 AspAspAspPheGluTrpAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeuAla 321
Db 824 GATGATGATTTTGAACCTAAGCTGTCATGACAGAAATTCGAGCTCTCTTAAAGT 883
Qy 322 ValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTrpAsnLysPro 341
Db 884 GTGAGCAAAATGCACATGAGCTTACCCAGATGAAAGACACATTTACTGGAGCATTCCT 943
Qy 342 GluProGlnProProTrpTrpAlaAlaSerAlaGlnPheArgArgAlaSerPheMetGly 361
Db 944 GCTGCTGCCGCCACATACATTCGACAGCTGCTGACTACTGATACCCCATTTCTGGGG 1003
Qy 362 Ser 362
Db 1004 TCA 1006

RESULT 3
US-09-768-826-16
; Sequence 16, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PF512P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1148)
```

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1209)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1229)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1266)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1285)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1287)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1290)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1295)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1305)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1324)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16

Alignment Scores:
Pred. No.: 5,89e-66 Length: 1350
Score: 662.00 Matches: 149
Percent Similarity: 61.06% Conservative: 36
Best Local Similarity: 49.17% Mismatches: 78
Query Match: 21.22% Indels: 41
Gaps: 9
US-09-622-964-3 (1-585) x US-09-768-826-16 (1-1350)
Qy 172 GlnLysLeuSerLeuProHisAsnMetPheTrpValITrPheAlaAsnLeu 191
Db 38 GAGAGCCTGAATCCGACTTCAACAGTACTGGGTCCCTGCTGCTGCTGCTGCTGCTG 97
Qy 192 SerMetLysAlaTrpLeuGlyArgIleArgAspProIleLeuGlnSerLeu 211
Db 98 GCGGCCAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 157
Qy 212 AsnGluMetAsnTrpLeuArgTrpGlnCysGlyHisLeuTrpAlaTrpIleSer 231
```

```
Db 158 GAAGAGCTGAACAAGTACGACGACCAAGTGCATGATATCCACTTGTGATGATACG 217
Oy 232 ILleProteinVal1YrThrGlnVal1Val1Val1Val1YrSerPhePheLeuThrCys 251
Db 218 ATCCCTCTCTCTACACCCCAAGTGTGATGACCATAGCCCTTACTTCTTCTTGGCCCTCC 277
Oy 252 LeuVal1GlyArgGlnPheLeuAsnPro-----AlaLysAlaTyr----- 254
Db 278 CTGGTTGGCCGCGCAAGTTTGTGGAGCCAGAGGCGGCGCTGCACAACTCAGAAAGCTTCG 337
Oy 265 ---ProGlyHisGlu-----LeuAspLeuVal1ProVal1Phe 276
Db 338 AAGCCAGCCAGGACGACGACCCGCTGGAGACCCGACACATGACGCTCTCCACC 397
Oy 277 ThePheLeuGlnPhePhePheTyrVal1GlyTyrLeuLysVal1AglGlnLeuIleAsn 296
Db 398 ACTCTGCTCAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 457
Oy 297 ProPheGlyGluAspAspAspAspPheGluThrAsnTyrPheVal1AspArgAsnLeuGln 316
Db 458 CCATTGGTGGAGATATGATGACGCTTGGAGACAAATCAGCTCTAGACGCGCAACTGGCAG 517
Oy 317 ValSerLeuLeuAlaVal1AspGluMetHisGlnAspLeuProArgMetGluProAspMet 336
Db 518 GTGTCCCTCTATCCCTGACGAAATGATACAGAACCTTCCCGCTCGAAGAGCAG 577
Oy 337 TyrTyrAsnLeuProGlnProGlnProGlnProGlnProGlnProGlnProGlnProGln 355
Db 578 TACTGGGATGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 637
Oy 356 ArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGlnGluMetGluPhe 375
Db 638 CGGCCCTCTATCTCTGGCTCCACCTTCAACCTGCGGATGACGAGCAGC----- 685
Oy 376 GlnProAsnGlnGlnLysPheGluLysAlaHisAlaGly----- 388
Db 686 ---CCGACGACGACGCTGACGAGGCGCTCCCGCGGATCTGTCGCGCGCGCGCGC 742
Oy 389 -----IleIleGlyArgPheLeuGlyLeu-----GlnSerHisAspHis 402
Db 743 GCGGACGACCCCTTCTGCTGCGCGCTTCTGCGCGCTGAGGGCGCCCTCCGCGCGC 801
Oy 403 ProProAlaGlnAsnSerArgThrLysLeuLeuTyrProLysArgLysSerLeuLeuHis 422
Db 802 CCTCCGGAAGTTCGCGCGCTGCGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCTCCG 861
Oy 423 GluGly-----LeuProLysAsn-----HisLysAlaAlaLysGlnAsnVal 436
Db 862 GCGGAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 921
Oy 437 ArgGlyGln 439
Db 922 CGGGGACGA 930

RESULT 4
US-10-027-632-134529/c
; Sequence 134529, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134529
; LENGTH: 748
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134529

Alignment Scores:
Pred. No.: 1,748-28 Length: 748
Score: 333.50 Matches: 83
Percent Similarity: 38.188 Conservative: 1
Best Local Similarity: 37.738 Mismatches: 6
Query Match: 10.698 Indels: 131
Gaps: 1
```

US-09-622-964-3 (1-585) x US-10-027-632-134529 (1-748)

```
Oy 152 ProSerAlaGlnHisLeuVal1GlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeu 171
Db 712 CCCACCGCTTTTCACTCCACTCCACTCCAGGCTTATGACTCCGCGACACACAGCAGTG- 654
Oy 172 GluLysLeuSerLeuProHisAsnMetPheTyrVal1ProTyrVal1TyrPheHisAsnLeu 191
Db 653 GAGAAACAGAGCTTACACACACACACACATGTTCTGTGGTCCCTGGGTGGTGGTGGTGGT 594
Oy 192 SerMetLysAlaTyrPheGluGlyArgGlyLeuArgAspProIleLeuLeuGlnSerLeuLeu 211
Db 593 TCAATGAAGCGCTGCTGAGTGCAGTAATCCGGAACCTATCTCTCCAGAGCCTGGCG 534
Oy 212 Asn----- 212
Db 533 AAGCTGAGCCCACTGTACAGACAGGCGTCCGCGACAGTGGGAAGGTTGTGTCACAG 474
Oy 212 ----- 212
Db 473 AAACAAGGTTTCTTCAAAAGAGAACCTTGGGCCCTGAGGCTTCCGAGAGCCGGAG 414
Oy 212 ----- 212
Db 413 TGGGTTGACAGAACTTTTCCACACAGCAATCCACAGCCGAGGTGCTCTTCAGAG 354
Oy 212 ----- 212
Db 353 CCCCTCCTCTTCTCCAAAGTCTGTGAGTCTCTGTTCTCTTTTATATAGAGAAAGCTG 294
Oy 212 ----- 212
Db 293 AGACAAAGAGGTTTACTAGAGCTTCCATGCGCACACAGCGAGGAATGACATAGTA 234
Oy 212 ----- 212
Db 233 CGAGGCTGTGTACTGTGAKAAGAGTGGGCGAGCCGAGGTGGGCGAGGTGGT 174
Oy 213 -----GluMetAsnThrLeuArgThrGlnGly 222
Db 173 CAGAACCCCAATCCCTCTTGTGCCCCCGACGAGATGAAACCTTGTCTACTCAGTGG 114
Oy 222 YHisLeuTyrAlaTyrAspTyrPheSerIleProLeuVal1TyrThrGlnVal1Thr 241
Db 113 ACACCTGATGCTTACAGACTGATAGTATGCCACTGTGTATACACAGGTGAGAGCT 56

RESULT 5
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11346
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
; OTHER INFORMATION: n = A,T,C or G
; US-10-198-846-11346

```

Alignment Scores:

```

Pred. No.: 5.48e-25 Length: 853
Score: 304.00 Matches: 56
Percent Similarity: 82.05% Conservative: 8
Best Local Similarity: 71.79% Mismatches: 14
Query Match: 9.74% Indels: 0
DB: 15 Gaps: 0

```

US-09-622-964-3 (1-585) x US-10-198-846-11346 (1-853)

```

QY 83 GlyPheTyValThrLeuValAlaThrArgTyrTrpAsnGlnTyrGluAsnLeuProTyr 102
DB 273 GGGTTTATGTTCTGCTGAGTGAACCGATGCGAACAGTTTGCAATTCGCCCTCG 332
QY 103 ProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGlnSerArg 122
DB 333 CCGAGACAGCTATGTTCCATCTCATCTGACGAGTTCACGAGCGACGCGCGCG 392
QY 123 LeuLeuArgArgThrLeuLeuLeuArgTyrAlaAsnLeuGlyAsnValLeuLeuLeuArgSer 142
DB 393 CGGCTTAGAAGGACGCTGATGCGCTACGCAATCTCACCTCCCTGCTCATCTTCCGCTCG 452
QY 143 ValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
DB 453 GTGAGCAGCTGCTGTACAAAGATTCCACAAATGACACACGCTGTTGACGA 506

```

RESULT 6

```

; US-10-027-632-134530/c
; Sequence 134530, Application US/10027632
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

```

```

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-134530

```

Alignment Scores:

```

Pred. No.: 8.35e-25 Length: 578
Score: 300.00 Matches: 56
Percent Similarity: 93.44% Conservative: 1
Best Local Similarity: 91.80% Mismatches: 2
Query Match: 9.62% Indels: 2
DB: 15 Gaps: 1

```

US-09-622-964-3 (1-585) x US-10-027-632-134530 (1-578)

```

QY 152 ProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeu 171
DB 201 CCTTCTCCTCCAC-----TCTGACAGCTTATGACTCCGCGAACAACAGCTTG 148
QY 172 GluLysLeuSerLeuProHisAsnMetPheTyrValProTyrValTyrPheAlaAsnLeu 191
DB 147 GAGAACTGAGCTTACACACAAACATGTTGGGTGCGCTGGCTGTTGCCACCTG 88
QY 192 SerMetLysAlaTyrPheLysGlyAlaTyrIleArgAspProIleLeuGlnSerLeuLeu 211
DB 87 TCATGAGGCGTGGCTTGGAGTGCATCCGAGACCTTATCTCTCCAGAGCTGCTG 28
QY 212 Asn 212
DB 27 AAC 25

```

RESULT 7

```

; US-10-198-846-2561/c
; Sequence 2561, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2561
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 390, 462
; OTHER INFORMATION: n = A,T,C or G
; US-10-198-846-2561

```

Alignment Scores:

```

Pred. No.: 8.2e-21 Length: 462
Score: 264.00 Matches: 49
Percent Similarity: 82.09% Conservative: 6
Best Local Similarity: 73.13% Mismatches: 12
Query Match: 8.46% Indels: 0
DB: 15 Gaps: 0

```

```
US-09-622-964-3 (1-585) x US-10-198-846-2561 (1-462)
QY 82 LeuGlyPheTyrValThrLeuValValThrArgTyrPaspnGlnTyrGluAsnLeuPro 101
DB 256 ATAGGGTTTATTTACTCTGCTGATGTAACCGATGGGAGACAGTTTGAAATTTGCC 197
QY 102 TrrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGlnGlnSer 121
DB 196 TGGCCAGACAGCTAATGTTCCATCTCAGACAGGTTTACCGAAGCGACGACGCGG 137
QY 122 ArgLeuLeuArgArgThrLeuLeuLeuArgTyrAlaAsnLeuGlyAsnValLeuLeuArg 141
DB 136 CGCTCTCTTAGAAGGAGCGGTGATGCGCTACGTCAATCTCACCTCTGCTCATCTTCCG 77
QY 142 SerValSerThrAlaValTyr 148
DB 76 TCGGTGAGCACTGCTGTGTAC 56

RESULT 8
US-10-198-846-7843
; Sequence 7843, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 7843
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-7843

Alignment Scores:
Pred. No.: 1,31e-20 Length: 615
Score: 264.00 Matches: 49
Percent Similarity: 82.09% Conservative: 6
Best Local Similarity: 73.13% Mismatches: 12
Query Match: 8,468 Indels: 0
DB: 15 Gaps: 0

US-09-622-964-3 (1-585) x US-10-198-846-7843 (1-615)
QY 82 LeuGlyPheTyrValThrLeuValValThrArgTyrPaspnGlnTyrGluAsnLeuPro 101
DB 352 ATAGGGTTTATTTACTCTGCTGATGTAACCGATGGGAGACAGTTTGAAATTTGCC 411
QY 102 TrrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGlnGlnSer 121
DB 412 TGGCCAGACAGCTAATGTTCCATCTCAGACAGGTTTACCGAAGCGACGACGCGG 471
QY 122 ArgLeuLeuArgArgThrLeuLeuLeuArgTyrAlaAsnLeuGlyAsnValLeuLeuArg 141
DB 472 CGCTCTCTTAGAAGGAGCGGTGATGCGCTACGTCAATCTCACCTCTGCTCATCTTCCG 531
QY 142 SerValSerThrAlaValTyr 148
DB 532 TCGGTGAGCACTGCTGTGTAC 552

RESULT 9
US-09-764-872-195
; Sequence 195, Application US/09764872
; Publication No. US20030050231A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL25
; CURRENT APPLICATION NUMBER: US/09/764,872
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 195
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (490)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-872-195

Alignment Scores:
Pred. No.: 9.54e-18 Length: 507
Score: 238.00 Matches: 46
Percent Similarity: 72.29% Conservative: 14
Best Local Similarity: 55.42% Mismatches: 23
Query Match: 7,638 Indels: 0
DB: 12 Gaps: 0

US-09-622-964-3 (1-585) x US-09-764-872-195 (1-507)
QY 1 MethrileThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 37 ATGAGGGTTTATTTACTCTCATAAGTGGCGGAGCCCTTCCGAGGTTCTCTGCGCTG 96
QY 21 LeuLeuCysTrrPArgLysSerLeuTyrLysLeuLeuTyrGlyGluPheLeuLeuLeu 40
DB 97 CTTCGCCCTGAGGAGGAGCATCTACAAGCTCTCTCAAGAGATTCCTCTTGGG 156
QY 41 LeuCysTyrTyrIleIleArgPheLeuTyrArgLeuAlaLeuThrGluGlnGlnLeu 60
DB 157 GCCTTGATGCTGATGCTTAGCATCACCCTGCGTGTGCTGACCCAGACAGAGGTAC 216
QY 61 MetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
DB 217 GTGATGCTCAGAGGAGGCCGCTACTGCAACCGCTCAGACAGACATTCCTTGTCTTT 276
QY 81 ValLeuGly 83
DB 277 GTATTGGGT 285

RESULT 10
US-10-198-846-1862
; Sequence 1862, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1862
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

NAME/KEY: misc_feature
LOCATION: 15, 553, 561, 570, 671, 734, 739, 745
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1862

Alignment Scores:

Pred. No.:	9,76e-10	Length:	751
Score:	171.00	Matches:	49
Percent Similarity:	55.05%	Conservative:	11
Best Local Similarity:	44.95%	Mismatches:	41
Query Match:	5.48%	Indels:	10
DB:	15	Gaps:	3

US-09-622-964-3 (1-585) x US-10-198-846-1862 (1-751)

QY 82 LeuGlyPheThyValThrLeuValValThrArgTTPTrpAsnGlnTryGluAsnLeuPro 101
DB 372 ATAGGGTTTATGCTACTGCTAGTGAACCCGATGGGAGACAGTTTGGAATTTGCC 431
QY 102 TTPProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGlnSer 121
DB 432 TG-CCAGACAGGCTAATGTCTCCTACATCTTACAGCTGTTCCACCGAAGCGACAGG 490
QY 122 ArgLeuLeuArgArgThrLeuLeuArgTyrAlaAsnLeuGlyAsnValLeuLeuArg 141
DB 491 CGCTGCTTAAAGAGCGCTGATGCCCTC-GTCAATCTACCTTCTGCTTACTTTCT 549
QY 142 SerValSerThAlaValTyrLysArgPhe-----ProSerAlaGlnHisLeu 157
DB 550 CGGAGCGCTGNTGTCTCGCCGCTTAACCTAGGAGCCCGGCTGCGAGATTC 609
QY 158 ValGlnAlaGlyPheMetThrProAla-----GluHisLysGlnLeuGlnLysLeuSer 175
DB 610 ATTAACCTATCATCCGTCATCTACGAGGCGCGCTCCACTTGTCTTAGTGGGTAANG 669
QY 176 LeuProHisAsn-----MetPheTrp 182
DB 670 CMCCTGGCACTGCATAGTGTCTGG 696

RESULT 11

US-10-027-632-174961/c

Sequence 174961, Application US/10027632

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

POLYMERISMS IN THE HUMAN GENOME

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

PRIOR FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 174961

LENGTH: 3186778

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(3186778)

OTHER INFORMATION: n = A,T,C or G

US-10-027-632-174961

Alignment Scores:

Pred. No.:	1.2e+03	Length:	3186778
Score:	118.00	Matches:	31
Percent Similarity:	68.52%	Conservative:	6
Best Local Similarity:	57.41%	Mismatches:	17
Query Match:	3.78%	Indels:	1
DB:	15	Gaps:	0

US-09-622-964-3 (1-585) x US-10-027-632-174961 (1-3186778)

QY 106 LeuMetSerLeuValSerGlyPheValGluGlyLysAspGlnSerArgLeuLeuArg 125
DB 285840 CTCCTAGACCTGCTGTCGCC-CTCCGCGAAGGAGGAGGAGGAGGAGGAGG 285782
QY 126 ArgThrLeuLeuArgTyrAlaAsnLeuGlyAsnValLeuLeuLeuArgSerValSerThr 145
DB 285781 AGCATGCTCTCCGCTCCGCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 285722
QY 146 AlaValTyrLysArgPheProSerAlaGlnHisLeuValGln 159
DB 285721 CTCGACACAGGCGCTTCCACAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 285680

RESULT 12

US-09-925-300-141

Sequence 141, Application US/09925300

Patent No. US20020151681A1

GENERAL INFORMATION:

APPLICANT: Craig Rosen,

APPLICANT: Steve Ruben,

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05988

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1890

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 141

LENGTH: 1630

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1566)

OTHER INFORMATION: n equals a,t,g, or c

US-09-925-300-141

Alignment Scores:

Pred. No.:	0.0115	Length:	1630
Score:	114.50	Matches:	102
Percent Similarity:	30.23%	Conservative:	41
Best Local Similarity:	21.56%	Mismatches:	149
Query Match:	3.67%	Indels:	184
DB:	11	Gaps:	20

US-09-622-964-3 (1-585) x US-09-925-300-141 (1-1630)

QY 77 ProIleSerPheValLeuGlyPheThyValThrLeuValValThrArgTTPTrpAsnGln 96
DB 225 CTTCTCTCTCTCCGCGGCGGCTGACGTCGTCGCTCACTTTAGAGTTTACTTCAACAC 284
QY 97 TyrGluAsnLeuProTrp-----Pro 103
DB 285 GTGAGAG--CTTCATGCGCGGCTCTCAGTCTGCGGAGAGATTAACATCTGTGG 341
QY 104 AspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp-----GluGlnSer 121
DB 342 GAGAGACTGTCAAGCTGGGAGGAGAC-CCGCTGGGAGACGACTGTCCGAGACAGAT 400

```

QY 122 ArgLeuLeuArgArgThrLeuLeuLeuArgTyrAlaAsnLeuGlnValLeuLeuArg 141
DB 401 AGGCTCCCCAGAGCTCTGG---AGGCAGAAAGTGTGCTCCACAGCTTGCGCCCTAGG 457
QY 142 SerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAlaGly 161
DB 458 -----CCGGA 463
QY 162 PheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsn----- 179
DB 464 GCTTCAGTGGCTCAGTCACACCGGTGGCCCTGGGACAGTCCCTTCCCTACAGATCCACAG 523
QY 180 MetPheTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyGly 199
DB 524 GTGTCTGTGATCCAGGCTCTGTGTG---CTGTGCGGTGGCTGTCTGTG 571
QY 200 ArgIleArgAspProLeuLeuLeuGlnSerLeuLeu-----AsnGluMetAsn 215
DB 572 CTGTGACGGGGCGGCTTAATTGGTCAACACTTACATGACTTTTCCAAAGGCGATTGACC 631
QY 216 ThrLeuArgThrGlnCysGlnHisLeuTyrAlaTyrAspTyr----- 229
DB 632 ACAGAAAAGAGTATGATGACAGGACACTGTGACCGAATCTTGGACCGGAGATGTCTGCC 691
QY 230 -----LleSerIleProLeuValTyrThrGlnVal 239
DB 692 GGTGGGAGCTCTCTACAGCTGGGCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 745
QY 240 ValThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsn 259
DB 746 -----ACCTGT----- 751
QY 260 ProAlaLysAlaTyrProGlnHisGluLeuAspLeuValValProValPheThrPheLeu 279
DB 752 -----CCCTCTGAAACGTGAGC 769
QY 280 GlnPhePhePheTyrValGlyTyrPheLysValAlaGlnLeuLeuLeuAsnProPheGly 299
DB 770 ACTGGCTC-----TTACTACTTTGGAGCGCTGTGCTGCTGCTGCTGCTGCT 814
QY 300 GluAspAspAspAspPheGluThrAsnTyrPheAlaAspArgAsnLeuGlnValSerLeu 319
DB 815 ACACAGAGGAATGATTCAGATTCAGTGGCTGTGGAGACCTTCATCTCCATCATCACT 874
QY 320 LeuAlaValAspGluMetHisGlnAspLeuProArgMet-----GluProAspMet 336
DB 875 TTGGCCCGCTGGCTGTGATGTTCGCCCTACGCCATCCAGGTGGGTCCTGGCCATCTGCC 934
QY 337 TyrTrpAsnLysProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArg 356
DB 935 CACTGTCTATGCCATCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
QY 357 AlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGlnGluMetGluPheGln 376
DB 985 -----TCCAACAACACCGGAGACATGGAGTC-CGA 1013
QY 377 ProAsnGlnGluAspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeu 396
DB 1014 CCGGGA-----GGCGG 1025
QY 397 GlnSerHisAspHisProProArgAlaAsnSerArgThrLysLeuLeuTyrProLys 416
DB 1026 TATGTGTACGCTGCCATCTCATCGGCGCCAC-----GTTCTCTCTA----- 1067
QY 417 ArgGlnSerLeuLeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnVal 436
DB 1068 ---CATCTCTACAAACACTGCTCTGCT----- 1094
QY 437 ArgGlyGlnGluAspAsnLysAlaTyrPylsLeuLysAlaValAlaPheLysSerGly 456
DB 1095 -----GCCCTA 1100
QY 457 ProLeuTyrGlnArgProGlyTyr-----TyrSerAlaProGlnThr-ProLeuSerPr 474

```

```

DB 1101 CTGTGCTTACGATCTGTGGCCACACACTGACACCATCAGCTGTGACTCCCTGCTTAC 1160
QY 474 othPromePhePheProLeuGlu-----Pr 483
DB 1161 CATTCGATGCGCTTCTCCCTTACAGACAGTTTCGAGGACGAGCCCTTCAACAAGTCC 1220
QY 483 oSerAlaProSerLysLeuHisSerValThrGlyIle 495
DB 1221 CCAGAGACTGCGCAACCTCAACCTCGTGGGACTT 1257

RESULT 13
US-10-171-581-275
; Sequence 275, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes In Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 275
; LENGTH: 3280
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: D80012
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-275

Alignment Scores:
Pred. No.: 0.0809 Length: 3280
Score: 111.50 Matches: 58
Percent Similarity: 34.27% Conservative: 27
Best Local Similarity: 23.39% Mismatches: 100
Query Match: 3.57% Indels: 63
DB: 15 Gaps: 10

US-09-622-964-3 (1-585) x US-10-171-581-275 (1-3280)
QY 337 TyrTrpAsnLysProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArg 356
DB 1460 TTCATAATATATGCCAGTACTCCAAACCCGACAAAGCTTTGAGATTAATATCGTAGG 1519
QY 357 AlaSerPheMetGlySerThrPheAsn-----LleSerLeuAsn 369
DB 1520 GATATTCGCCCTGGAGCTGCTTGAAGCCACATATATTAACAGACTCTGCACATTAC 1579
QY 370 LysGlnGluMetGluPheGlnProAsnGlnGluAspGluGluAspAlaHisAlaGlyIle 389
DB 1580 AAGTCAGAGCGTGTGAAAAGGTGCGACAAAGATGCTGAGAA----- 1624
QY 390 IleGlyArgPheLeuGlyLeuGlnSerHisAspHisProProArgAlaAsnSerArg 409
DB 1625 -----TACTTAGGCTTCATTTCTAAATGACTTCTAGAGAAATGTGAACTAAG 1675
QY 410 ThrLysLeuLeuTyrProLysArgGlnSerLeuLeuHisGlnGlyLeuProLysAsnHis 429
DB 1676 ---AAGCTTCTCTACCAAGATTAATGAAAACCTTTCAGATTCCAAAGGCCCAAAAACAC 1732
QY 430 -----LysAlaAlaLysGlnAsnValArgGlyGlnGluAsp----- 441
DB 1733 TCGGTCAATGAAGAAGCAGGAAGAACAGGTGAAGAACGAGATGGGAACAA 1792
QY 442 -----AsnLysAlaTyrPylsLeuLysAlaValAlaPheLysSerGlyPro 457
DB 1793 GTGGGCCCGCGAACAAGACTTCCTGCACCCGCGCGCGAT---TTTGTTCACAGCTCA 1849

```

OY 458 Leu-----TyrGlnArgProGlyTyrTyr 465
DB 1850 ATCACCAGCATTTTGGTGACACATCAGCTGTGCTGTTACACAGAGTTCAAAAGAA 1909
OY 466 SerAlaProGlnThrProLeuSerProThrPrometPhePheProLeuGluProSerAla 485
DB 1910 TCTGCCACTTTGCAGCCA-----TTTTCACGTTGCAG----- 1942
OY 486 ProSerLysLeuHisSerValThrGlyLeuAspThrLysAspLysSerLeuLysThrVal 505
DB 1943 -----TTGGATATCCAGTCAGACAGATAGTACGACAGTC 1975
OY 506 SerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAlaProGlyAlaLeuMetGlu 525
DB 1976 CAGATGACACTGAGAGCTGTGTCGCAAGACATCTGTCAGAGTTATACCAAAACC 2035
OY 526 HisProGluValSerGlnValArgArgLysThrValGluPheAsnLeuThrAspMetPro 545
DB 2036 AAACMAAGAGTTGAGATTAAGTCGAAGAGTACTGTGAA-----AAACTCCCT 2083
OY 546 GluIleProGluAsnHisLeuLys 553
DB 2084 CCTGTCTCTGCTGCTGCACCTGAAA 2107

RESULT 14
US-10-114-893-87
Sequence 87, Application US/10114893
Publication No. US20020193567A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKeough
APPLICANT: Kelleher, Kerry S.
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10/114,893
EARLIER FILING DATE: 2002-04-02
EARLIER APPLICATION NUMBER: 09/413,232
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 87
LENGTH: 1276
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-893-87

Alignment Scores:
Pred. No.: 0.0195 Length: 1276
Score: 111.00 Matches: 64
Percent Similarity: 37.79% Conservative: 35
Best Local Similarity: 24.43% Mismatches: 97
Query Match: 3.56% Indels: 66
DB: 15 Gaps: 10

US-09-622-964-3 (1-585) x US-10-114-893-87 (1-1276)

OY 341 ProGluProGlnProProTyrThrAlaAla-----SerAlaGlnPheArg 355
DB 328 CCAAGAAACCAACCCATGCTACTGTCAGACAGATCATGTGATTCAGAGTCCAGAA 387
OY 356 ArgAlaSerPheMetGlySerThrPheAsnLeu-----SerLeuAsnLysGlu 371
DB 388 AGCGATCCGACATTCACAGCACAGAAAAATTATTCAGACCCCTCAAGTCTACAAAGCCA 447

OY 372 GluMetGluPheGlnProAsnGlnGluAspGluGluAspAlaHisAlaGlyIleIleGly 391
DB 448 TCCATA----- 453
OY 392 ArgPheLeuGlyLeuGlnSerHisAspHisHisProProAlaAlaAsnSerArgThrLys 411
DB 454 -----CCAAACACTGCAGAAAGTTAACTAAT 480
OY 412 LeuLeuTrpProLysArg-----GluSerLeuLeuHisGlnGluLysLeuProLysAsnHis 429
DB 481 CCATCATATTCAGAAAAAGCTATCCAGATCATGCTCAAAAAAATTAAAGCAAGTCACT 540
OY 430 LysAlaAlaLysGlnAsnValArgGlyGlnGluAsp-----AsnLysAla 444
DB 541 CACCTGGAAAAAGCACAATAGAAAGGACGACCTAGAAAATATATAGCTAAGCTATAGCG 600
OY 445 TrpLysLeuLysAlaValAlaPheLysSerGlyProLeuTyrGlnArgProGlyTyr 464
DB 601 TGTAAAGCTACGACATTCAGATTAAGCCAGTCAGGCCACCTCAGCTATTCAAGCCACTTAT 660
OY 465 TyrSerAlaProGlnThrProLeuSerProThrPrometPhePheProLeuGluProSer 484
DB 661 TCATCTCATCCAGAAATGAATCTCACATCCAGCCATTGCTCCACAGAAATTGGCT 720
OY 485 AlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLysSerLeuLysThr 504
DB 721 AACCTCCCAACATTTTAAT-----CCAAAAAGTCA 753
OY 505 ValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAspGlyAlaLeuMet 524
DB 754 GTGAGTCTAGGACGGCA-----GCCTTATTACCAACTGTGAATTAGT----- 798
OY 525 GluHisProGluValSerGlnValArgArgLysThrValGluPheAsnLeuThrAspMet 544
DB 799 -----GAACCTGTCAACCCCTACAGAAAAAATGTTTGGCAAAACTTTATAGC 849
OY 545 ProGluIleProGluAsnHisLeuLysGluProLeuGlnGlnSerProThrAsnIleHis 564
DB 850 CCTTTGGTC-----AATGATATTCTGAGGCAAGAGAAAAAACTCAAAACCTAAT 903
OY 565 ThrThr-----LeuLysAspHisMetAspProTyrThrAlaLeuLysAsnArgAspGlu 582
DB 904 GTTTCAGCAAGAAAGTCATCTCTCCAGGCTCCTTGTGAATGATTCCAGGAACAT 963
OY 583 AlaHis 584
DB 964 GCATAC 969

RESULT 15
US-10-027-632-245664/c
Sequence 245664, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
EARLIER FILING DATE: 2002-04-30
EARLIER APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720

```

; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 245664
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-245664

```

```

Alignment Scores:
Pred. No.: 0.00659
Score: 110.50
Percent Similarity: 64.58%
Best Local Similarity: 50.00%
Query Match: 3.54%
DB: 15
Length: 609
Matches: 24
Conservative: 7
Mismatch: 12
Indels: 5
Gaps: 1

```

US-09-622-964-3 (1-585) x US-10-027-632-245664 (1-609)

```

QY 43 TYRTYRILELLEArg-----PheileTYRArgLeuAlaLeuThrGlu 57
DB 409 TATATATCATTAAGGCGTGCATTTTATTTTGTCTCCAGATTGTACTTACAGAGTC 350
QY 58 GlnGlnLeuMetPheGlnLysLeuThrLeuTYRCysAspSerTYRileGlnLeuIlePro 77
DB 349 CAAATAACGTTACTTGAATAATTATCAATTACTGTGACAGATATGCTGAACAAATTCCA 290
QY 78 IleSerPheValLeuGlyPheTYR 85
DB 289 GTAACCTTGTGCTGTGTAAGTAT 266

```

Search completed: July 27, 2003, 20:48:30
Job time : 1278.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:08:55 ; Search time 72.2406 Seconds

(without alignments)
2483.449 Million cell updates/sec

Title: US-09-622-964-3
Perfect score: 3120
Sequence: 1 MITTYSQVAMNRLGFSRL.....TLKHDPYWALENDEAHNS 565

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09622964/runat.22072003_101119_26036/app.query.fasta_1.2065
-DB=Issued_Patents_NA -OFMT=fastp -SUFFIX=rm1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09622964.GC.N1.1.79.6runat.22072003_101119_26036 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEBOUENR -NEG_SCORES=0 -WAIT -OSPBLOCK=100 -LONGLOG
-DEV_TIMECUT=120 -WARN_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued_Patents_NA:*
2: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/PCUTS.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	3.5	4152	4	US-08-826-134-5
2	108	3.5	5294	4	US-08-826-134-1
3	107.5	3.4	1202	3	US-08-728-603-16
4	107.5	3.4	32207	2	US-08-770-379-20
5	107.5	3.4	32207	4	US-08-757-669A-20
6	107.5	3.4	32207	4	US-09-230-371A-20
7	100	3.2	4143	4	US-08-826-134-6
8	100	3.2	5350	4	US-08-826-134-3
9	100	3.2	6585	4	US-08-746-111-4
c 10	98.5	3.2	2121	1	US-08-574-043A-1
c 11	98.5	3.2	2121	1	US-08-795-015-1
c 12	98.5	3.2	2121	4	US-08-456-297-1

c 13	98.5	3.2	2121	5	PCT-US94-12936-1	Sequence 1, Appl1
c 14	98.5	3.2	2121	5	US-08-272-255-5	Sequence 1, Appl1
c 15	98.5	3.2	2458	5	US-08-272-255-5	Sequence 5, Appl1
c 16	98.5	3.2	2458	5	PCT-US95-08565-5	Sequence 5, Appl1
c 17	98.5	3.1	8758	4	US-09-134-001C-1492	Sequence 1492, Ap
c 18	98.5	3.1	2363	4	US-08-487-596-5	Sequence 3, Appl1
c 19	97.5	3.1	3545	4	US-08-885-291-54	Sequence 54, Appl1
c 20	97.5	3.1	3545	4	US-08-486-672-54	Sequence 54, Appl1
c 21	97.5	3.1	5715	3	US-09-107-847-1	Sequence 1, Appl1
c 22	97.5	3.1	3645	2	US-08-663-112-1	Sequence 3, Appl1
c 23	96.5	3.1	6048	4	US-09-634-920-3	Sequence 2122, Ap
c 24	96.5	3.1	2613	4	US-09-134-001C-2122	Sequence 21, Appl1
c 25	95.5	3.1	3435	4	US-09-046-158A-21	Sequence 17, Appl1
c 26	95.5	3.1	1977	1	US-08-272-255-17	Sequence 17, Appl1
c 27	94.5	3.0	1977	5	PCT-US95-08565-17	Sequence 42, Appl1
c 28	94.5	3.0	3854	2	US-08-599-455B-42	Sequence 42, Appl1
c 29	94.5	3.0	3854	4	US-09-069-781B-42	Sequence 42, Appl1
c 30	94.5	3.0	3854	4	US-09-137-132-42	Sequence 42, Appl1
c 31	94.5	3.0	3854	4	US-08-664-564A-42	Sequence 42, Appl1
c 32	94.5	3.0	3854	4	US-09-094-410-42	Sequence 42, Appl1
c 33	94.5	3.0	802	2	US-09-018-595B-2	Sequence 2, Appl1
c 34	94	3.0	802	3	US-09-324-709A-2	Sequence 12, Appl1
c 35	94	3.0	43676	3	US-08-658-340-1	Sequence 1, Appl1
c 36	94	3.0	6909	2	US-08-804-196-1	Sequence 26, Appl1
c 37	93.5	3.0	6909	3	US-08-567-508C-1	Sequence 1, Appl1
c 38	93.5	3.0	6909	3	US-08-746-111-26	Sequence 1, Appl1
c 39	93.5	3.0	4482	2	US-09-196-480-1	Sequence 1, Appl1
c 40	92.5	3.0	4482	2	US-08-359-850-3	Sequence 219, App
c 41	92.5	3.0	1002	2	US-08-961-083-219	Sequence 5, Appl1
c 42	92	2.9	1879	3	US-08-961-527-5	Sequence 14, Appl1
c 43	92	2.9	13926	4	US-08-765-907A-14	
c 44	92	2.9	2220	4		
c 45	90.5	2.9				

ALIGNMENTS

RESULT 1
US-08-826-134-5
; Sequence 5, Application US/08826134A
; Patent No. 6465210
; GENERAL INFORMATION:
; APPLICANT: Pelas, Elliot
; TITLE OF INVENTION: CASPR/190, A FUNCTIONAL LIGAND FOR RPTP-BETA AND THE
; FILE REFERENCE: 7683-111
; CURRENT APPLICATION NUMBER: US/08/826,134A
; CURRENT FILING DATE: 1997-03-26
; EARLIER APPLICATION NUMBER: 60/014,199
; EARLIER FILING DATE: 1996-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-826-134-5

Alignment Scores:
Pred. No.: 0.135
Score: 108.00
Percent Similarity: 31.56%
Best Local Similarity: 20.53%
Query Match: 3.46%
DB: 4
Gaps: 22

US-09-622-964-3 (1-585), x US-08-826-134-5 (1-4152)

OY 142 serValserThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAlaIacy 161
DB 2010 ACAGTGCATGACAGTTCCTGCTGACATATCCCGGCTGCTCAACACTGCAGAGGCTACCC 2069
OY 162 PhenethrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsnMetPhe 181

```

Db 2070 CTACAGCTTTGGATTGGCCGAATGAGAGCAGACACTTCTACTGGGAGAGCTCCAGCC 2129
QY 182 TTPVALPROTPVALTPRPHLEAASLEUSERMETLYSALATPRLEUGLYGATGILE 201
Db 2130 TGGGATCCCA-----GGCCTGTGCTGTGGTGGTCCAGCGAGCTG 2168
QY 202 ATGASPRIOILEULEUENGLINSEULEUENGLIUMETASNTPLHLEUATGTHGLINDY 221
Db 2169 TGTGACAGCTCCCTTACTGACAACTGTGACCGCTGACAGCCCACTGAGAACTGACA 2228
QY 221 GGLYHLSLEUATYRATYRASPTRIPLESERILEPEOLEUVALTYRTHGLINVAL-- 240
Db 2229 GGGACTGCTGACCTTGTGGACCACTGCTGTC-----ACTCAGGTACTGAT 2276
QY 241 -----ThValAlaValTyrSerPhePhe-----LeuThCysLe 252
Db 2277 AGGGATACGAACCGCTCCACTCTGAGCCCAAGTTCTTCTGAGCGCTGCTGCTGTA 2336
QY 252 VALGLYATGGLNPRHLEUASNPFOALALYSALATYRPROGLY----- 266
Db 2337 TGGCGATCGAATTCCTGGAACACCATTCCTTCACACCGGGGGTGCACCTAGCTCC 2396
QY 267 -----HisGluLeuAspLeuValValProValPheThrPheLeuGlnPh 281
Db 2397 CCCAATCCGTGCCAACCACAGCTGGAT-----GTCTCCTT 2432
QY 281 ePhePhe-----TyrVal 285
Db 2433 CTACTTCAGGAGACCTGTGCTCCCTGGGGGTCTTCTTAAGAAATGAGGGGCCCTTACTG 2492
QY 285 LGLYTP-----LeuLysValAlaGluGlnLeuLeuAsnPr 297
Db 2493 CCAGTGGCGCGACCTTATGTGCGGGGTGAACCTCAACACATCCCGGATGTGTCTGCG 2552
QY 297 oPhe-----GlyGluAspAsp-----AspAspPheGluThAs 308
Db 2553 CTTGATGTGGGAGATGGGAGATGAGAACCTCACTACACTCACTGACGACTTGAATTCAA 2612
QY 308 n-----TrpIleValAspArg-----AsnLeuGlnValSerLeuLeuAlaVal 322
Db 2613 TGATGACAGATGGACCTGTGCGGGGTGAATACGTGAAGACGAGCGCCGCTCCGAGT 2672
QY 322 LASPGLUMETHISGLINSPLEUPROARGMETGLUPROASPMETTYRTPASNLSPROGL 342
Db 2673 GGATCACCGCGCCCTGGGTTCTGCGGCTATGTCACCTGACGACCTACATCTGATGAGATA 2732
QY 342 UPROGLNPROPTOTYRTHRAALASERIALAGLNPHLEARGTARGLASERPHMETGLYSE 362
Db 2733 TGACCAAGCCCTCAT---GTGGGATCTGCGAGACCTTAAGAGAGCCCTTTGTGGGTG 2789
QY 362 rThPheAsnLLeSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnLysAspG1 382
Db 2790 CTTGAGGGCCATGCTGTGAACGAGTACTCTG-----AACCTGAGAGGGCCG 2837
QY 382 uGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAspHis-- 401
Db 2838 TGGCAATGCCCTGTGAGGTACTCTACCACTGACAGGCGACATGTGCTCCACCTCGGCT 2897
QY 402 -----HisProArgAlaAsnSerArgThrLysLeuLeuThrProLysArgG1 418
Db 2898 CCCCTGTTTCCATGAGAGCGCGCTGCGAGCGGTATAGCTACTACACGCTGTGACTGTGA 2957
QY 418 uSerLeuLeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgG1 438
Db 2958 CCTACAGCGCTTGTGATGGCCATCTGCAACAC---GATATGTGTG 3002
QY 438 yGlnGluAspAsnLysAlaTrpLys----- 446
Db 3003 TTTCTTTGAGACCGGGACCTGTGATGGCGCTTAACTACACTGACGCGCTGCTGACGC 3062
QY 447 -----LeuLysAlaValAspAlaPheLysSerGly-----Pr 457

```

```

Db 3063 CAGGAGTTCTCCACATCTGAGCGCGGACCTGCCAGCTTATGAGCCTTGACTCC 3122
QY 457 oLeuTyrglnArgProGlyTyr-----TyrSerAlaProG1 469
Db 3123 GGGGTATGATPACTCCGGGCTATGTGCTGCTGCTGCTACATGAGCCCGGCTAGCGCTGCCGA 3182
QY 469 nThrProLeuSerProThrProMetPhePheProLeuGluProSerAlaProSerLysLe 489
Db 3183 CTACCCCGCGCTGCTGCTGCTG-----CCCGGTACCGTGGCCTGT 3227
QY 489 uHisSerValThrGly-----IleAspThrLysAspLysLe 501
Db 3228 CTACAAAGTTAGGGAGAGAGAGTCTCTCTACGTTCTGACACACACACTCCGCGCTGCT 3287
QY 501 rLeuLysThrValSerSerGlyAlaLysSerPheGluLeuLeuSerGlnSerAspG1 521
Db 3288 CCTGCTACAGTCACTCTTGTGCTGAGTACATGAGCTGTGCTGCTGCTGCTGCTGCTGCT 3347
QY 521 yAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsnLe 541
Db 3348 GACCTTACG----- 3371
QY 541 uThrAspMetProGluLeuProGluAsnHisLeuLysGluProLeuGluGlnSerProTh 561
Db 3372 GGGCAGCAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
QY 561 rAsnLLeHisThrThr 566
Db 3432 TAGCATCAATATCAC 3447

RESULT 2
US-08-826-134-1
; Sequence 1, Application US/08826134A
; Patent No. 6465210
; GENERAL INFORMATION:
; APPLICANT: Peles, Elior
; TITLE OF INVENTION: CASPR/130, A FUNCTIONAL LIGAND FOR RPRP-BETA AND THE
; FILE REFERENCE: 7683-111
; CURRENT APPLICATION NUMBER: US/08/826,134A
; EARLIER FILING DATE: 1997-03-26
; EARLIER APPLICATION NUMBER: 60/014,199
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 5294
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218)..(4369)
US-08-826-134-1

Alignment Scores:
Pred. No.: 0.208 Length: 5294
Score: 108.00 Matches: 108
Percent Similarity: 31.56% Conservative: 58
Best Local Similarity: 20.53% Mismatches: 213
Query Match: 3.46% Indels: 147
Gaps: 22

US-09-622-964-3 (1-585) x US-08-826-134-1 (1-5294)
QY 142 SerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAlaGly 161
Db 2227 ACAGTGAATCGAGTCTCTCTACCAATTCCTCCGCTGCTCAACACATGAGAGGCTACCC 2286
QY 162 PheMetThrProAlaGlnHisLysGlnLeuGlyLysLeuSerLeuProHisAsnMetPhe 181
Db 2287 CTACAGCTTTGGATTGGCCGAATGAGAGACAGCACTTCTTACTGTGGGAGGCTCCAGCC 2346
QY 182 TTPVALPROTPVALTPRPHLEAASLEUSERMETLYSALATPRLEUGLYGATGILE 201

```

```

Db 2347 TGGATCA-----GCGCTGCTGCTGCTGCTGACCGGAGCTG 2385
QY 202 ArgAspProIleLeuLeuGlnSerLeuLeuAsn-GluMetAsnThrLeuArgThrGlnCy 221
Db 2386 TGTGGACCCGCTTGTACTGCACTGTGACGCTGACACAGCCGACAGTGGAGAACTGTGACA 2445
QY 221 sGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValVal 240
Db 2446 GGGACTGCTACCTTTGTGGACCATGTGCTGTC-----ACTGAGTAAGTAT 2493
QY 241 -----ThrValAlaValTyrSerPhePhe-----LeuThrCysLe 252
Db 2494 AGGGAGATACGACCGGCTGCACTTCTGAGCGCCAGTTCCTGCTGAGCGCTGTGGCTGCTA 2553
QY 252 uValGlyArgGlnPheLeuAsnProAlaTyrAlaTyrProGly----- 266
Db 2554 TGGCGATCGAAATTCCTGGAACACATTCCTTCACACCGGGGCTGCTACGCTTCC 2613
QY 267 -----HisGluLeuAspLeuValValProValPheThrPheLeuGlnP 281
Db 2614 CCCAATCCGTCGCAACACACGCTGAT-----GTCCTCTT 2649
QY 281 ePhePhe----- 285
Db 2650 CTACTTCAGGACCTGTCTCCCTCGGGGCTTCTCTAGAAATATGGGGGCGCTTACTG 2709
QY 285 lGlyTrp-----LeuLysValAlaGluGlnLeuLeuAsnPr 297
Db 2710 CCAGTGGCGCCGACCTTATGTGGGGGTGAGACTCACACATCCCGGAGTGTGCTTCC 2769
QY 297 oPhe-----GlyGluAspAsp-----AspAspPheGluThrAs 308
Db 2770 CTGTGATGTGGGATGGGATGGGATGAGAACTCACAGTACACTACAGAGACTTGTGAGTTCAA 2829
QY 308 n-----TrpIleValAspArg-----AsnLeuGlnValSerLeuLeuAla 322
Db 2830 TGATGACGAGTGGACCTGTGCTGGGCTGAATCAACGTAAGAGAGAGCGGCTCGAGCT 2889
QY 322 lAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLysProG 342
Db 2890 GGATACCGGCGCTGCTGCTGAGCGGAGTACTG-----AACCTGGAGGCGC 3054
QY 342 uProGlnProTyrThrAlaAlaSerAlaGlnPheArgAlaGlnPheMetGlySe 362
Db 2950 TGACACGCGCCCTAT-----GTGGGATGTGCAAGACTTAAGAGACCCCTTGTGGGCTG 3006
QY 362 rThrPheAsnIleSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnLysProG 382
Db 3007 CTGAGGGCGCATGCTGTGAACGGAGTACTGT-----MACCTGGAGGCGC 3054
QY 382 uGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlnSerHisAspHis-- 401
Db 3095 TGGCAATCCCTGTGAGGTACTCAACCCACTGCACAGGCACTGTGCCACCCCTCGGCT 3114
QY 402 -----HisProAlaGlnAlaAsnSerArgThrLysLeuLeuTyrProLysArgG 418
Db 3115 CCCCTGTTCCATGAGGCGCGCTGCGTGAAGCGCTATAGCTACTACAGCTGTGACTGTGA 3174
QY 418 uSerLeuLeuHisGluGluLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgG 438
Db 3175 CCTACAGGCTTTTATGTGGCATACTGCACAC-----GATATTGGTGG 3219
QY 438 yGlnGluAspAsnLysAlaTrpLys----- 446
Db 3220 TTTCTTTGAGCGCGGACACTGGATGCCCTATACCTACAGTACAGCGCTGCGCTGTGAGC 3279
QY 447 -----LeuLysAlaValAspAlaPheLysSerGly-----Pr 457
Db 3280 CAGGAGTCTTCCACATGTGTGAGCGGCGAGTCCAGGCTATAGCGCTGCTACATCC 3339
QY 457 oLeuTyrGlnArgProGlyTyr-----TyrSerAlaProG 469

```

```

Db 3340 GGGCTATGATACCTCGGGCTATGTGCTGCTGCTACCATAGGCCCGGGTACCGGCTGCCGA 3399
QY 469 nThrProLeuSerProThrProMetPhePheProLeuGluProSerAlaProSerLysLe 489
Db 3400 CTACCCCGCGCGCTGCTGCGCTGTG-----CCGGTTACGGTGGCGCTGT 3444
QY 489 uHisSerValThrGly-----IleAspThrLysAspLysLe 501
Db 3445 CTACACGTTACAGGAGAGAGAGTCTCTTACGCTTACAGACAGCAGCTCCCGCTGCTGT 3504
QY 501 rLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerLysAspG 521
Db 3505 CCGCTCTACGCTAGTCTGCTTGTGTGCTGACTACATAGCGCTGTCTCTCAAGATGATGG 3564
QY 521 yAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsnLe 541
Db 3565 GACCCCTGAG-----CTCGGATATGAGCT 3588
QY 541 uThrAspMetProGluIleProGluAsnHisLeuLysGluProLeuGlnGlnSerProTh 561
Db 3589 GGGCACCAGTCCCTACGTGTACACCTAACCACTGACAGTACCGATGGCCAGCCCA 3648
QY 561 rAsnIleHisThrThr 566
Db 3649 TAGCATCAATATACACC 3664

```

RESULT 3

US-08-728-603-16
Sequence 16, Application US/08728603
Patent No. 6093806

GENERAL INFORMATION:

APPLICANT: Cesarman, Ethel
APPLICANT: Knowles, Daniel M.
TITLE OF INVENTION: PROTEINS OF KAPOSI'S SARCOMA ASSOCIATED
NUMBER OF INVENTIONS: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728, 603
FILING DATE: 10-OCT-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: BRAMAN, SUSAN J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 19603/720
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1636
TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1202 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
US-08-728-603-16

Alignment Scores:

Pred. No.: 0.0172
Score: 107.50
Percent Similarity: 38.25%
Best Local Similarity: 24.40%
Length: 1202
Matches: 81
Conservative: 46
Mismatch: 117

Query Match: 3.45% Indels: 89
DB: 3 Gaps: 17

US-09-622-964-3 (1-585) x US-08-728-603-16 (1-1202)

QY 3 IIEThrYThSerInValAlaAsnAlaArgLeuGlySerPheSerArgLeuLeu 22
DB GTTCACTTATTTTTCAGACAGCCGATCGCGGAGAGCATATATACGCTCCG 307
QY 23 -----CysTrpArgLysSerIleTyrLysLeu---LeuTyrGlyLupheLeu 39
DB GGTATGCTTAACCTCGCTGTCTTACCATATCTCTATGGCACAAGCTGTGATG 367
QY 40 LeuLeuGlySerIleLeuArgPheIleTyrArgLeuAlaLeuThrGluGln 59
DB TTGTTTCCCATATATATCTCCACAGGCTGTGCAGCTT-----GAA 409
QY 60 LeuMetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSer 79
DB ATTTTCTTACTATTTATGTCTACTGATATCTTC---AGTGTGTGCGCTCACT 466
QY 80 PheValLeuGlyPheTyrValThrLeuValValThrArgTrpAsnGlnTyrGluAsn 99
DB CTAGTC---AGTACCTCTGTGGCATAATCTACGCGCTTCTGCGCCACAAGAGAGTCC 523
QY 100 LeuProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluLysAsp--- 118
DB CTGCGATGGGACTGACATCCGCTGACATGTAATTCATGTGTGCTGCGGGGATGCC 583
QY 119 -----GluGlnSerArgLeu-----LeuAlaArgThrLeuIleArgTyrAlaAsn 133
DB TGTCCACACAGAGAGAGGCTGTGCAGCCGCTGACAGCAGCCCATGTTATGAGAAC 643
QY 134 LeuGlyAsnVal-----LeuIleLeuArgSerValSerThrAlaValTyr 148
DB GCGGGAACATGACACGACAGCTGGGAGCTGATGTCAGAACCGTGTCA----- 691
QY 149 LysArgPheProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHis 168
DB GGTCTGCTGTTTTTTTGTGTTTGTCTTCCCTTACACAGTAACTAATCTAGGACACTCT 873
QY 169 LysGlnLeuGluLysLeuSerLeuProHisAsnMetPheTrpValTrpProTyrValTrp 188
DB TGTGCGCTCTTATCTGTTTATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
QY 188 AlaAsnLeuSerMetLysAlaTrpLeuGlyLysArgIleArgAspProIleLeuLeuGln 208
DB AGGAGACAAAGCTGCAGCAAGCA-----GCGGGAAGTA-AGGGGGGATGTTGCTGT 813
QY 208 nSerLeu-----LeuAsnGluMetAsnThrLe 217
DB GGTCTGCTGTTTTTTTGTGTTTGTCTTCCCTTACACAGTAACTAATCTAGGACACTCT 873
QY 217 u-----ArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIle 231
DB GCTAAGGCGAGCGTGCATCCGCGGACAGCTGC-----TATACGCGGGGTTGATAA 924
QY 231 rIleProLeuValTyrThr-----GlnValValThrValAlaValTyrSerPhe 248
DB CGTGGGCTGCGACATCTGTTACTGACAGCACTGACAGCCCGCGTGTCCCGAT 984
QY 248 eLeuThrCysLeu-----LeuValGln 254
DB ATACTCTGCTGCGATCCCTCTTACGACAGAGATGACGCTCTTCCAAAGCTCAG 1044
QY 254 yArgGlnPheLeuAsn-----ProAlaLysAlaTyrProGlnHisGlu 269
DB GCACTCTTTCATGTCGCGGCGACACAGAGTACCGCGGATGTCTACGCTTCCCT 1104
QY 269 uAspLeuValValProValPheThrPheLeuGln 280
DB TAAATTAATCTAGCTCCGCTTCCCAAGATGACAG 1138

RESULT 4

US-08-770-379-20

Sequence 20, Application US/08770379

Patent No. 5849564

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Edelman, Isidore S.

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED

TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/770, 379

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28, 678

REFERENCE/DOCKET NUMBER: 52342

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 32207 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-770-379-20

Alignment Scores:

Pred. No.: 5.85

Score: 107.50

Percent Similarity: 38.25%

Best Local Similarity: 24.40%

Query Match: 3.45%

DB: 2 Gaps: 17

US-09-622-964-3 (1-585) x US-08-770-379-20 (1-32207)

QY 3 IIEThrYThSerInValAlaAsnAlaArgLeuGlySerPheSerArgLeuLeu 22

DB GTTCACTTATTTTTCAGACAGCCGATCGCGGAGAGCATATATACGCTCCG 24331

QY 23 -----CysTrpArgLysSerIleTyrLysLeu---LeuTyrGlyLupheLeu 39

DB GGTATGCTTAACCTCGCTGTCTTACCATATCTCTATGGCACAAGCTGTGATG 24391

QY 40 LeuLeuGlySerIleLeuArgPheIleTyrArgLeuAlaLeuThrGluGln 59

DB TTGTTTCCCATATATATCTCCACAGGCTGTGCAGCTT-----GAA 24433

QY 60 LeuMetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSer 79

DB ATTTTCTTACTATTTATGTCTACTGATATCTTC---AGTGTGTGCGCTCACT 24490

QY 80 PheValLeuGlyPheTyrValThrLeuValValThrArgTrpTrpAsnGlnTyrGluAsn 99

DB CTAGTC---AGTACCTCTGTGGCATAATCTACGCGCTTCTGCGCCACAAGAGTCC 24490

Db 24785 AGGAGGACCAAGCTGCAAGCCA-----GGGGGAAGGA-AGGGGGCTGATGTGCTGT 24837
 Oy 208 nsearleu-----Leuansglumelanthrile 217
 Db 24838 GGTGCTGCTGTTTTTTGTGTTTGTCTTCCTTACCACGTAATCTAATCTAGCACTGT 24897
 Oy 217 U-----ArgThnGlnCysGlyHisLeuTyraLeuTyraSerPhe 231
 Db 24898 GCTAAGGCGAGCGCTGATCCGGGACAGCTGC-----TATACGGGGGGGTGATPAA 24948
 Oy 231 rllrleuValLeuValTythr-----GlnValValThrValAlaValTyraSerPhe 248
 Db 24949 CCGTGGCTGCGACAGTAACCTCGCTTACTGACGAGGCACTGTACAGCGCGGTGTCCCTGAT 25008
 Oy 248 eleuThrCysLeu-----Valgl 254
 Db 25009 ATATCCCTGCGCTGGGATCCCTCTTTAGGACAGAGGATACGGTCTCTCCAAAGCCTCAG 25068
 Oy 254 yArgGlnPheLeuAsn-----ProAlaValAlaTyraProGlyHisGluLe 269
 Db 25069 GCAGCTTTCATGTCGCGGCGCCACACGTAAGCCGCGAGTGTACTAGTGCCCTCCCT 25128
 Oy 269 uaspleuValValProValPheThrPheLeuGln 280
 Db 25129 TAATTATATGAGCTCCCGTCCCATGATGACG 25162

RESULT 6
 US-09-230-371A-20
 ; Sequence 20, Application US/09230371A
 ; Patent No. 6348586
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Yuan
 ; APPLICANT: Bohenzky, Roy A
 ; APPLICANT: Russo, James J
 ; APPLICANT: Edelman, Isidore S
 ; APPLICANT: Moore, Patrick S
 ; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
 ; FILE REFERENCE: 45185-G-PCT-US
 ; CURRENT APPLICATION NUMBER: US/09/230, 371A
 ; CURRENT FILING DATE: 1999-11-17
 ; PRIOR APPLICATION NUMBER: PCT/US97/13346
 ; PRIOR FILING DATE: 1997-07-22
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 32207
 ; TYPE: DNA
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus
 US-09-230-371A-20

Alignment Scores:

Pred. NO.:	5.85	Length:	32207.
Score:	107.50	Matches:	81
Percent Similarity:	38.25%	Conservative:	46
Best Local Similarity:	24.40%	Mismatches:	117
Query Match:	3.45%	Indels:	89
DB:	4	Gaps:	17

US-09-622-964-3 (1-585) x US-09-230-371A-20 (1-32207)

Oy 3 lIleThrTyThrSerGlnValAlaAsnAlaArGleuGlySerPheSerArgLeuLeuLeu 22
 Db 24272 GTCACTACATATTTTGGCAAGCACCGATCGCGGGAGGAGCGATGATGATATCTGCTCTG 24331
 Oy 23 -----CysTrpArgGlySerIleTyraValLeu-----LeuTyraGlyGlnPheLeuIlePhe 39
 Db 24332 GGTATGTGCTTAACCTCGCTGTGCTTACCAATATCTATTGGCAGAAAGTGTGATGTTT 24391
 Oy 40 LeuLeuCysTyrrTyrrIleIleIleArPheIleTyraArgLeuAlaLeuThrGlnGlnGln 59
 Db 24392 TTGTTCCTCCAAATATATCATCTCCACAGGCTTGTCACACTT-----GAA 24433

60 Leuethheglutylserthrleutyrcysaspertryllelnleproleser 79
Db 24434 ATTTTCTTACTATTATAGTGCTACTGCGATATCTC---AGGTGGTGCCGCACST 24490

80 PheValLeuGllyPheTyrrValThrlleuValValThrArgTrprAsnGlnTygIasn 99
Db 24491 CTAGTG---AGGTACCCTTCGGTGGCATATCTCAATCCGCTTCCTGCCAAGAAGACATCC 24547

100 leuprotrpraasparglemeSerleuValSerGlyPheValGluGlyAsp--- 118
Db 24548 CTGGGATGGGTACTGACATCCGCTGACCTGTAATTCGATTTGGCTGTGCGGGGATGCC 24607

119 -----GlnInserArGLEU-----LeuargrThrleuLleargTyrlaasn 133
Db 24608 TCTGCACACAGCAGACGAGGTGTGTCGACCCCGGTGACAAAGACGCATGTGTTATGAGAAC 24667

134 LeuGlyAsnVal-----LeuLleuArgSerValSerThrAlaValTyr 148
Db 24668 GCGGGGAACATGACCTGCAGACTGGCGACCTGCATGCATGAAACCGTGTC- 24715

149 LysArgPrheProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHis 168
Db 24716 -----GTTACTGCAGGTTTTCCGTTATCCC----- 24739

169 LysGlnLeuGluLysLeuSerleuProHisAsnMetHetrPrvalProTrp-ValTrph 188
Db 24740 -----CTGGCCCTCCCTTAATCTGTATTATGCTCACCTGCTGTCTGTGTCG 24784

188 ealaAsnLeuSerMetLysAlaTrpLeuGlyGlyArgLileArgAspProLleLeuLeuG 208
Db 24785 AGGAGACAAGCTGACAGCCA-----GGCGGAAGTAA-AGGGGGGTGATTGTTCGCT 24837

208 nSerleu-----LeuasnlunetAsnThrLe 217
Db 24838 GGTCGTGCTGTTTTTTGTGTGTTTGCCTCCCTACACAGTCAATAATCTAGTGGACACTCT 24897

217 u-----ArgThrGlnCysGlyHisLeuTyrrAlaTyrAspTrpIleSe 231
Db 24898 GCCTAAGCGACGCTGGATCCGGACAGCTCC-----TATACGGGGGGGTGATTA 24948

231 rIleProLeuValTyrThr-----GlnValValThrValAlaValTyrSerPhePh 248
Db 24949 CGTGGCTCTGGCAGTAGTACCTCGTTACTCGACGACACTGTACAGCGCCGTGGTCCCTGAT 25008

248 eLeuThrCysLeu-----ValGI 254
Db 25009 ATACTCTCTCTGGGATCCCTCTTTAGGACAGAGATGATCGCTCTTCCAAAGCCCTGAG 25068

254 YArgGlnPheLeuAsn-----ProAlaLysAlaTyrProGlyHisGluLe 269
Db 25069 GCAGCTCTTTCATGTCCGGCGCCACACAGTAGCCCCGGGAGTGTACGTGCCCTCCCT 25128

269 uAspLeuValProValPheThrPheLeuGln 280
Db 25129 TAATTTAATCTAGCTCCCGCTTCCCATGATGACAG 25162

US-08-826-134-6
; Sequence 6, Application US/08826134A
; Patent No. 6465210
; GENERAL INFORMATION:
; APPLICANT: Pellex, Eljor
; TITLE OF INVENTION: CASPR/190, A FUNCTIONAL LIGAND FOR RPTP-BETA AND THE
; FILE REFERENCE: AXONAL CELL RECOGNITION MOLECULE CONTRACTIN
; CURRENT APPLICATION NUMBER: US/08/826,134A
; CURRENT FILING DATE: 1997-03-26
; EARLIER APPLICATION NUMBER: 60/014,199
; EARLIER FILING DATE: 1996-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6

```

      .....g.....m.....c.....n.....e.....r.....o.....y.....s.....e.....l.....i.....n.....f.....h.....e.....s.....t.....l.....e.....u.....s.....n.....l.....y.....s.....c.....l.....
||||:: ::||| |||::||| :: |||||

```

QY	Db	2760	TCCAGAGCTTAAGAGAGGCCCAATTTGTGGGTGCTTGAGGSSCAATGCGTTGAATGAGT	2819
QY	371	UGUmetGluPheGlnProAsnGlnGluAspGluAspAlaHisAlaGluLeuIleG1	391	
Db	2820	GACTGTG-----ACCTTGGAGGGTGTGTCACATGCTCCGAGGCACTTCCC	2867	
QY	391	YArgHeuGluGluGlnSerHisAspHis-----HisProAlaGlnAs	407	
Db	2868	CAATGCAAGGGGCACTGCACCCACCCCGGTTCCCTGTTCACAGGAGGACGTGTGT	2927	
QY	407	nSerArgThrIlyLeuLeuThrProIlyArgGluSerLeuLeuHisGluGluProIy	427	
Db	2928	GGAGGGATACAGACTACTACAGTGTGACGTGTAACSTCACAGCTTTTGATGAGCAATATTG	2987	
QY	427	SAsnHisIlySalaIalaIyGlnAsnValArgIyGlnGluAspAsnIlyAlaArgIlyS--	446	
Db	2988	TAATTCAC-----GATATGTGTGATTTCTTGAGACTGGCAATGATGGG	3032	
QY	447	-----LeuIy	448	
Db	3033	CTATAACCTCCAGTCAGACACTGCTGTGCGGCCAGAGATTCTGCATGCTGAGCCG	3092	
QY	448	sAlaValAspAlaPheIySerGly-----ProLeuArgIlnArgProGlyIlyr-----	464	
Db	3093	GCCGGTACCGGGGTATAGCGCTGGGTATATCCAGGCTACAGACACTCCGTGGTACCTGCC	3152	
QY	465	-----TyrSerAlaProGlnThrProLeuSerProThrPromerPh	478	
Db	3153	TGGGTACCATGTCCTGGGTGCGCGCTACCCGACACTACCAAGGCTGGCGCGCAAGTG--	3210	
QY	478	ePheProLeuGluProSerAlaProSerIlyLeuHisSerValThrGly-----	494	
Db	3211	-----CCGGATACCGGGGGCCCGGTGACATGTACTGTGAGAGAGAGCTTC	3257	
QY	495	-----IleAspThrIlyAspIlySerLeuIlyThrValSerSerGlyAlaIly	510	
Db	3258	CTTTAGCTTCAGACACAGCTGCTGCTCGACGTGCTGTGATGACGTGACACTCTTTGTGCG	3317	
QY	510	sIySerPheGluLeuLeuSerGluSerAspGlyAlaLeuMet-----	524	
Db	3318	TGACTACATGCGCGTGTCTATCAAGAAGATGGGACCTTCAGCTTCGCTATCAGCTGGG	3377	
QY	525	GluHisProGluValSerGlnValArgArgIlyThrVal	537	
Db	3378	CACCAATGCCCTATGTGTACAGACGATACACACCCGGCCAGTG	3417	
RESULT 8				
US-08-826-134-3				
Sequence 3, Application US/08826134A				
Patent No. 6465210				
GENERAL INFORMATION:				
APPLICANT: Peles, Eljor				
TITLE OF INVENTION: CASPR/190, A FUNCTIONAL LIGAND FOR RPTP-BETA AND THE				
FILE REFERENCE: 7683-111				
CURRENT APPLICATION NUMBER: US/08/826,134A				
EARLIER FILING DATE: 1997-03-26				
EARLIER APPLICATION NUMBER: 60/014,199				
NUMBER OF SEQ ID NOS: 32				
SOFTWARE: PatentIn Ver. 2.0				
SEQ ID NO 3				
LENGTH: 5350				
TYPE: DNA				
ORGANISM: Rattus norvegicus				
FEATURE:				
NAME/KEY: CDS				
LOCATION: (154)..(4296)				
US-08-826-134-3				
Alignment Scores:				
1.82 Length: 5350				
Score: 100.00				

Percent Similarity: 31.32% Conservative: 53
 Best Local Similarity: 21.01% Mismatches: 190
 Query Match: 3.21% Indels: 163
 DB: 4 Gaps: 26

US-09-622-964-3 (1-585) x US-08-826-134-3 (1-5350)

```

OY 142 SerValSerThrAlaValIleuValArgProSerAlaGlnHisLeuValGlnAlaGly 161
DB 2166 GCAGTGGATGAGATCTCTCCGTACACATTCGGCGCTCTCAACAC----- 2210
OY 162 PheMetThrProAlaGlnHisLysGlnLeuGluLysLeuSerLeuPro-----HisAsn 179
DB 2211 -----TCAGAGAGCTAACCCCTACAGCTTTTGGATTGGCCGAATGAAGAAC 2258
OY 180 MetPhe-----TrrValProTrrValTrrPheAlaAsnLeuSer 192
DB 2259 GCATTCTCTCTGGGAGGCTCCCGAGCTGGGATCCA-----ACG 2297
OY 193 MetLysAlaTrrPheGluGlyArgIleArg-AspProIleLeuLeuGlnSerLeuLys 212
DB 2298 CTGTGCTGCTGGGCTGGACCAAGCTGTATAGACCTTGCACCTGCACCTGCACCTGCATGC 2357
OY 212 nGluMetAsnTrrLeuArgTrrGlnCysGlnHisLysTrrAlaTyrAspTrrIleSerIle 232
DB 2358 TCACACGCCACAGTGGAGAACAGACAAAGGGGCTCTGACCTTTGTGGACCATCTGCTCTG 2417
OY 232 eProLeuValIleTrrGlnValIle-----ThrValAlaValIleTyrSe 246
DB 2418 C-----ACTCAGGTAGTGTATAGGTAGACACAAACCCCTCAGCTGTGAAGCTCA 2465
OY 246 trPhePhe-----LeuThrCysLeuValAlaGlyArgGlnPheLeuAsn----- 259
DB 2466 GTTCTTCCGAGGCTCTGCTGCTTATGTGTACGCGCATTCCTGGACACATATCTCTT 2525
OY 260 -----ProAlaLysAlaTyrProGlnHisGlnLeuAs 270
DB 2526 CCGCAGCTGAGCTGCAGCTGCTTCCCTCCCAATCCGCTGCC-----AACCAAGCAGCTTGA 2579
OY 270 PheuValValProValPheThrPheLeuGlnPhePhePhe----- 283
DB 2580 T-----GTCTCCTTCTACTTCAAGACCTGGCTCCCTCAGG 2615
OY 284 -----TrrValGlyTrr----- 287
DB 2616 AGTCTTCTAGAACATGGGGGCTCTTCTGCCAGTGGGCGCAGCTTACGTAGACT 2675
OY 288 -----LeuLysValAlaGlnGlnLeuIleAsnProPhe-----GlyGluAspAs 302
DB 2676 GGAGCTCAACACATCCCGGATGTGTGTTTGGCTTGTATATTGGCAATGGGAGATGAGAA 2735
OY 302 P-----AspAspPheGlnTrrAsn-----TrrIleValAspArg-- 313
DB 2736 CCTGACAGTGCACATGACTTGCAGTTCATGATGATGACAGAGTGGGCTTGGTCCGGGC 2795
OY 314 -----AsnLeuGlnValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProAr 331
DB 2796 TGAATATCAACCTGAGAGCGCCGCTGCGAGTGCAGACATCGGCGCTGGGTGCTAAAGCC 2855
OY 331 gMetGluProAspMetTrrTrrPasnLysProGluProGlnProProTrrAlaAlaIle 351
DB 2856 CATGCCCTGCAGACGATCATCTGCTGGAGTATGACCAACCCCTCAT--GTGGGATC 2912
OY 351 rAlaGlnPheArgTrrAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysG 371
DB 2913 TGCAGAGCTTAAAGGGGCCCATTTTGGGCTGCTTGAAGGCCATGCGCTTTGAATGAGAGT 2972
OY 371 uGluMetGluPheGlnProAsnGlnGluAspGluGlnAspAlaHisAlaGlyIleIleG 391
DB 2973 GACTCTG-----AACTTGGAGGGGTCTGCCAATGCTCCCGAGGACACTTCCC 3020
OY 391 yArgPheLeuGlyLeuGlnSerHisAspHis-----HisProProAlaGlnAs 407
  
```

```

DB 3021 CAATGACAGGGGCACTGCACCCACCCCGTTCCCTGTTTCCAGSAGAGAGCTGTGT 3080
OY 407 nSerArgThrLysLeuLeuTrrProLysArgIleuSerLeuHisGlnGlyLeuProLys 427
DB 3081 GAGGCGATACAGCTACTACAGCTGTGACTGTGTACACAGCTTTTGAAGSACCATATTGG 3140
OY 427 sAsnHisLysAlaAlaValGlnAsnValArgGlnGlnGlnAspLysValAlaTrrLys-- 446
DB 3141 TATATCAC-----GATATTGTGGTGGATTCTTGAAGACGGACATGGATGGG 3185
OY 447 -----LeuLys 448
DB 3186 CTATACCTTCACGTACAGCTGCTTGGGCGCCAGAGTTCTGTACATGCTGTAGCCG 3245
OY 448 sAlaValAlaAspAlaPheLysSerGly-----ProLeuTrrGlnAlaArgProGlyTrr-- 464
DB 3246 GCCGGTACCGGGGCTATGAGCCCTGGCTATATCCAGCTACAGCACTCTGGTATGCTGCC 3305
OY 465 -----TrrSerAlaProGlnTrrProLeuSerProThrProMetPhe 478
DB 3306 TGGGTACCATGGCTCTGGGTACCGCTACCCAGTACCCAAAGGCTGGCGGCGCAGTG-- 3363
OY 478 ePheProLeuGluProSerAlaProSerLysLeuHisSerValThrGly----- 494
DB 3364 -----CCCGATACCGGGGCGCGCTGTACATGTATTCTGAGAGAGAGCTGTC 3410
OY 495 -----IleAspThrLysAspLysSerLeuLysTrrValSerGlyAlaLys 510
DB 3411 CTTTAGCTTACAGACACAGCTGTGCTGTGACGTCTGTCTACAGTCACTCTTGTGTGG 3470
OY 510 LysSerPheGlnLeuLeuSerGlnSerLysAspGlyAlaLeuMet----- 524
DB 3471 TGACTACATGGCGGCTGTATCATTCAAGAGATGGGACCCCTACAGCTTCCGTATACAGCTGG 3530
OY 525 -GluHisProGlnValSerGlnValArgLysThrVal 537
DB 3531 CACAGTCCCTATGTGTACACAGTACACACCCGCGCAGTG 3570

```

RESULT 9
 US-08-746-111-4
 ; Sequence 4, Application US/08746111
 ; Patent No. 6066778
 ; GENERAL INFORMATION:
 ; APPLICANT: Ginsburg, David
 ; TITLE OF INVENTION: Compositions And Methods For Screening
 ; TITLE OF INVENTION: Compounds For Anticoagulant Activity
 ; NUMBER OF SEQUENCES: 54
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/746,111
 ; FILING DATE: 06-NOV-1996
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ingolia, Diane E.
 ; REGISTRATION NUMBER: 40,027
 ; REFERENCE/DOCKET NUMBER: UM-02536
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:

LENGTH: 6585 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 6..6554
US-08-746-111-4

Alignment Scores:
Pred. No.: 2.64 Length: 6585
Score: 100.00 Matches: 85
Percent Similarity: 34.388 Conservative: 57
Best Local Similarity: 20.588 Mismatches: 147
Query Match: 3.21% Indels: 124
Gaps: 19

US-09-622-964-3 (1-585) x US-08-746-111-4 (1-6585)

```
QY 216 ThleuargThrcincysglYhlsleu-----TyrLatyrAsprTrpIleSerIlePro 233
DB 2598 ACCATGACAGCAGGAGACCCACATGATGACAGCAGGTTCTCTGATGAAGGCCCA 2657
QY 234 LeuValTyTrhGlnValValThrValAlaValTySerPhePheLeuThrcysLeuVal 253
DB 2658 GCCTGCTAAACT----- 2669
QY 254 GlyArgGlnPheLeuAsnProAlaLysAlaTyProGlyHlsGlnLeuAsnLeuValVal 273
DB 2670 GGGAGGCAT--TCAAAACCAAGAAATGCTGATTCGAAATGAGCTGAGGAGCAT 2726
QY 274 Pro-----ValPheTrhPheLeuGlnPhePhe 283
DB 2727 CCTAGGAGTGTATACCTTAAAGCAAAAGATCACTCCAAATTTCTGAATGAGCATGG 2786
QY 284 TyrValGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyLysAspAsp 303
DB 2787 CCGTGGGCTTCGAAAAGGCTGATGTAATTAATTAAGCAAAATGTCAGACACAGAT 2846
QY 304 ---AspPheGlnThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeuAlaVal 322
DB 2847 GTGGATAGCTGACCAACAGCTCTCAAAATCATGAATATACAGTA----- 2891
QY 323 AspGlnMetHlsGlnAsnLeuProArgMetGlnProAspMetTyTrpAsn----- 339
DB 2892 -----CCTGGGGAGAGAGACACCTCTCACACAAACACACACAGA 2930
QY 340 LysProGlnProGlnProProGlyTrhAlaAlaSerAlaGlnPheArgAlaSerPhe 359
DB 2931 AAGCCAAAGTGAACCTCCAAACATTTCTGAGCTGGACATAA----- 2972
QY 360 MetGlySerTrhPheAsnIleSerLeuAsnLysGlnGlnMetGlnPheGlnProAsnGln 379
DB 2973 -----TCTCCACATGTAAAGACAGGAGAAACAGTGGTTTTCAGAAAAGACAG 3023
QY 380 -----GlnAspGlnGlnAspAlaHlsAlaGlyLe 389
DB 3024 TTATTCATCAGACAGCAAGAAAGAAATAAGAAAGCTTGCACTACAGATCCCTCTA 3083
QY 390 IleGlyArg---PheLeuGlyLeuGlnSerHlsAsnHlsPro-----ProArg 405
DB 3084 TCTCCAAAGGGCTTTGACCTTTGAGAGGACATAAACCTCCCATTTCCAGACAGGAGA 3143
QY 406 AlaAsnSerArgTrhLysLeuLeuTrpLysArgGlnSerLeuLeuHlsGlnGlyLeu 425
DB 3144 CTACTTAATCACTACCTGTTACTCCACAAGTCCAAATGACGCTCTTCCAGACCTG 3203
QY 426 ProLysAsnHlsLysAlaAlaLysGlnAsnValArgGlnGlnLysAspAlaValTrp 445
DB 3204 AACCAAGACCTCTCTCAATGATGAGGACAGCTCACTCTCTGACTATTAATCAAGTACTCG 3263
```

```
QY 446 LysLeuLysAlaValAlaAspAlaPheLysSerGlyProLeuTyrglnAspProGlyTyTrp 465
DB 3264 AAAATATCACTGACGACGATGAGCTCTTCTTACATCTTATACG----- 3308
QY 466 SerAlaProGlnTrhProLeuSerProThProMetPhePheLeuGlnProSerAla 485
DB 3309 TCAGTGGCCGACAGGAAACACTCTCCAAACA-----TTCTCTGCCCAAGAT----- 3353
QY 486 ProSerLysLeuHlsSerValThr----- 493
DB 3354 CCGATCAAAACACACTTACACAGACAGATCTAGCTACAGATCCTCTGCCGAGACTCAGC 3413
QY 494 ---GlyIleAsp----- 496
DB 3414 CAGGGCTTATATGACCTTACATGACTTTTACCTGATGACATTTGGTCTAACATCT 3473
QY 497 ---ThrLysAspLysSerLeuLysTrh-----ValSer 506
DB 3474 TCTTTCAGACCAAGTCAAAAGTCATCTTCTTCCAGATGATGACCAAGCAATCCCT 3533
QY 507 SerGlyAlaLysLysSerPheGlnLeuLeuSerGlnSerAspGlyAlaLeuMetGlnHls 526
DB 3534 TCTCAGACTTACGCTCTTACATCTCTCCAGAAATGATGATGACAAATATTAT--TAC 3590
QY 527 ProGlnValSerGlnVal-----ArgArgLysTrhValGlnPheAsn 540
DB 3591 CCAGACCTGATGATGATGCTCTCTTCTTCCAGAAACATCAATGACCTCTCCGACAG 3650
QY 541 LeuThrAspMetProGlnLeuProGlnAsnHlsLeuLys 553
DB 3651 CTGGGCCAGGTGCCCCCTTCTCAGATGACAAACAGAG 3689
```

RESULT 10

US-08-574-043A-1/C
Sequence 1, Application US/08574043A
Patent No. 5807692
GENERAL INFORMATION:
APPLICANT: Kinzler, Kenneth W.
APPLICANT: El-Deliry, Wafik
APPLICANT: Vogelstein, Bert
TITLE OF INVENTION: p21WAF1 Derivatives and Diagnostic
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LMD
STREET: 1001 G Street, NW suite 1100
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/574,043A
FILING DATE:
CLASSIFICATION: 51A
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107,49698
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202,508,9209
TELEFAX: 202,508,9209
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```

? HYPOHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? CELL LINE: GM
? POSITION IN GENOME:
? MAP POSITION: 6p21.2
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 76..568
US-08-574-043A-1

Alignment Scores:
Pred. No.: 0.53      Length: 2121
Score: 98.50      Matches: 71
Percent Similarity: 29.81%      Conservative: 22
Best Local Similarity: 22.76%      Mismatches: 112
Query Match: 3.16%      Indels: 107
DB: 1      Gaps: 15

US-09-622-964-3 (1-585) x US-08-574-043A-1 (1-2121)
OY 263 ALATYRPGLYHISGLULEUASPHEUVALPROVALPHEPTH-PHELEUGLNPHEPHE 282
DB 1793 GCATGCCCTGCTCATAGCCTC----- 1773
OY 283 PHEITYRVALGILYRPLEULYSVALAAGLUGLNULE-----ASN 296
DB 1772 ---TACTGCCACCATCTTAAATGTCTGCTCTTCCGCTGCTCAATCAAGTGCAT 1716
OY 297 PROPHIEGLYGLUASPASPASPHEGLUTHRASTPILVALASPARGASLEUGLN 316
DB 1715 GAACGTGGGAGGAGGAGGTGATGAGGAGCGCTGACAT---TGAGGGGCCAGT 1659
OY 317 VALSERLEUVALVALASPLUETHISGLASPHEUVALPHEUVALPHEUVALPHE 1659
DB 1658 GTCCTCCCTC-----CTAGAAAGATCTACTCCCATCA 1626
OY 337 TYRTPANLSPROGLUPROGLINPROPTRYTHRALALASERLAGLNPHE----- 354
DB 1625 TATACCCCTAACACAGATATAACCCCACTCAAGGGGGCTGTGCCACACATGGACCT 1566
OY 355 -----ARGARGALASERHEMETGLYSERTHPHEASNLLESERLEUASN 369
DB 1565 CACCCCCACACTAGAGGAGGGGGCGGTGCTG-----CTTGAG 1527
OY 370 LYSGLUGLUMETGLUPHE----- 375
DB 1526 CTGCCTGAGTAGAAGCTAGGCTGCTCTTCTGTGTGCTCCCTCCCTCCAGTCCAT 1467
OY 376 -----GLNPROASGLNGLUASPGLUGLNUASPRLAHLAAGLYLLE 389
DB 1466 TGACGTGGGGGTGGGGGTGGGACAGGACCTCAGAGCCATCGAGCATGAGGGTACT 1407
OY 390 ILEGLYARPHLEUGLYLEUGLINSERHISASPHLSHPROPRDHALASNSERARG 409
DB 1406 GAAGGGAAGAACAGAGGGGAGGACAGCAGAGGGAATTGCAAGCCAGCTGGAGA 1347
OY 410 THRILYSLEULEUTP-----PROLYSARGLUSERLEUEN 421
DB 1346 AGAAGGGTAGAGCTGGGCTCTCAAAAGTACAGGGAGGCCAAAGAGGAAAAGCTCAA 1287
OY 422 HIS---GLUGLYLEUPROLYSASNLHISLYSALAIALYSGLINSVALARGLYGLN 440
DB 1286 CACTGACAGAGGGGCTCC-----AGTGCAGGTGAG-----AGGGGCCATGAG 1245
OY 441 ASPANLYSALATRPLYSLEULYSALVALASPLAHPHELYSSERGLYPROLEU----- 458
DB 1244 GGCAGGGGGGTGTCTGCTCAG-----TACCCACACTTCCCTGCTCTCA 1197
OY 459 -----TYRGLNARGPROGLYTYRTRYSERALAPROGLINHPROLEU--- 472
DB 1196 CAAGACAGAGGGGGGTATCAAGACCGAGGAGGTCCACAGTCCACACTCAGGCTAGGC 1137

```

```

OY 473 -----SERPROTHRPROMETPHE 478
DB 1136 TGTGCTCACTTCAGGGTCAACCTGCCCAACCTTAGAGAGGTGAGGGGACTCCAAAGTTT 1077
OY 479 PHEPROLEUGLUPRO-----SERALAPROSERLYSLEULHISERVALTHRGLYLEASP 496
DB 1076 TTGATGATGATCCCCCACTCGGTAGAGGCCCTTCAAAGTCCCATCTGTACTCTCAAAATG 1017
OY 497 THRILYSAPLYSERLEULYSERLTHVALSERSERGLY 508
DB 1016 AAAAGCATTCAGGTCTGAGTGTCCAGCAAGAGGGG 981

RESULT 11
US-08-795-015-1/C
? Sequence 1, Application US/08795015
? Patent No. 5871968
? GENERAL INFORMATION:
? APPLICANT: Kinzler, Kenneth W.
? APPLICANT: El-Deiry, Malik
? TITLE OF INVENTION: p21WAF1 Derivatives and Diagnostic
? NUMBER OF SEQUENCES: 7
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Banner & Allgrettl, LTD
? STREET: 1001 G Street, NW suite 1100
? CITY: Washington
? STATE: DC
? COUNTRY: USA
? ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/795,015
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/574,043
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Kagan, Sarah A. 32,141
? REGISTRATION NUMBER: 01107,49698
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202,508,9100
? TELEFAX: 202,508,9299
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2121 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Homo sapiens
? CELL LINE: GM
? POSITION IN GENOME:
? MAP POSITION: 6p21.2
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 76..568
US-08-795-015-1

Alignment Scores:
Pred. No.: 0.53      Length: 2121
Score: 98.50      Matches: 71
Percent Similarity: 29.81%      Conservative: 22

```

Best Local Similarity: 22.76% Mismatches: 112
 Query Match: 3.16% Indels: 107
 DB: 2 Gaps: 15

US-09-622-964-3 (1-585) x US-08-795-015-1 (1-2121)

```

OY 263 AlATyrProGlyHisGluLeuAspLeuValAlProValPheThrPheLeuGlnPhePhe 282
DB 1793 GCATGCCCTGCCATAGCCCTC-----
OY 283 PheTyrValGlyTyrPLeuLysValAlaGluGlnLeuLe-----Asn 296
DB 1772 ---TACTGCCACCATCTTAATAATGCTGACCTCTTCCCTGCTAATCAAGTGCAT 1716
OY 297 ProPheGlyGluAspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGln 316
DB 1715 GAATCGGGAGGGATGGGTGGATGAGAGAGAGTGCCTGAGAGATT---TGAGGGGCCAGT 1659
OY 317 ValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMet 336
DB 1658 GTCCTCCCTC-----CTAGAAAGATCTACTCCCTCCCATCA 1626
OY 337 TyrTrpAsnLysProGluProGluProTyrThrAlaAlaSerAlaGlnPhe----- 354
DB 1625 TATACCCCTTAACACAGAGATTAACCCCTCAAGGGGGCTGTCGCACACATGGAGACCT 1566
OY 355 -----ArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsn 369
DB 1565 CACCCCAACAGCAGTAAGAGAGGGGGGCTCCCTG-----CTTGAG 1527
OY 370 LysGluGluMetGluPhe----- 375
DB 1526 CTGCCTGAGGTAGACTAGAGGTGCTCTTCTTGTGTGTCCTTCCCTCCAGTCCAT 1467
OY 376 -----GlnProAsnGlnGluAspGluGlnAspAlaHisAlaGlyIle 389
DB 1466 TGAGCTGGGGGTGGGGGTGGGACAGCACCCTCAGACCCACCTGACCATGAGAGGTACT 1407
OY 390 IleGlyArgPheLeuGlyLeuGlnSerHisAspHisHisProProArgAlaAsnSerArg 409
DB 1406 GAAGGGAAGACAAAGAGGGAGGACAGCAGCAGAGGGGAATTGCAGACCCAGCTGGA 1347
OY 410 ThrLysLeuLeuTrp-----ProLysArgLysSerLeuLeu 421
DB 1346 AGAAGGGTAGCTGGGGCTCTCAAAAGGTACAGGGGACCAAGAGGAAAAAGCTCA 1287
OY 422 His---GluGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
DB 1286 CACTGAGAGGGGCTCCCTC-----AGTGCAGGTGAG-----AGGGGCATGAG 1245
OY 441 AspAsnLysAlaTrpLysLeuLysAlaValAlaPheLysSerGlyProLeu----- 458
DB 1244 GGCAGGGGGGTGTGCTGCTCAG-----TACCCACCTTCCCTGCTCTCA 1197
OY 459 -----TyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeu--- 472
DB 1196 CAAAGCAGAGGGGGGTATCAAGAGCAGAGGGGTACCAAGTCCAGCTCAGCCTAGGC 1137
OY 473 -----SerProThrProMetHe 478
DB 1136 TGTGCTCACTTCAGGGGTACCCCTGCCCAACCTTACAGAGAGTGAAGGAGGACCTCA 1077
OY 479 PheProLeuGluPro-----SerAlaProSerLysLeuHisSerValThrGlyIleAsp 496
DB 1076 TTGATGAGAGCCCACTGCTGAGAGGCCCTTCAAAAGTCCATCTGTTACTTCAATG 1017
OY 497 ThrLysAspLysSerLeuLysThrValSerSerGly 508
DB 1016 AAAAAGAAATTCAGGTCTGAGTGTCCAGAGAAAGGGGG 981

```

RESULT 12
 US-08-456-297-1/c
 ; Sequence 1, Application US/08456297

```

; Patent No. 6300059
; GENERAL INFORMATION:
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth W.
; TITLE OF INVENTION: Tumor Suppressor WAF1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW suite 1100
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,297
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/149,829
; FILING DATE: 10-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.44722
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: GM
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 6p21.2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 76..568
; OTHER INFORMATION: /gene="WAF1"
; US-08-456-297-1

Alignment Scores:
Pred. No.: 0.53 Length: 2121
Score: 98.50 Matches: 71
Percent Similarity: 29.81% Conservative: 22
Best Local Similarity: 22.76% Mismatches: 112
Query Match: 3.16% Indels: 107
DB: 4 Gaps: 15

US-09-622-964-3 (1-585) x US-08-456-297-1 (1-2121)
OY 263 AlATyrProGlyHisGluLeuAspLeuValAlProValPheThrPheLeuGlnPhePhe 282
DB 1793 GCATGCCCTGCCATAGCCCTC-----
OY 283 PheTyrValGlyTyrPLeuLysValAlaGluGlnLeuLe-----Asn 296
DB 1772 ---TACTGCCACCATCTTAATAATGCTGACCTCTTCCCTGCTAATCAAGTGCAT 1716
OY 297 ProPheGlyGluAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGln 316

```

```

Db 1715 GAAGTGGGAGGATGGGATGAGGAGAGTGGCTGAGCATTT---TGAGGGGCCACT 1659
Oy 317 ValserleuLeuAlaValaspGluMetHisGlnaspLeuProArgMetGluProAspMet 336
Db 1658 GTCTCCCTC-----CTAGAAAGATCTACTACCCCATCA 1626
Oy 337 TytTPrasnLysProGluProGluProTyrtThrAlaLaserAlaGlnPhe----- 354
Db 1625 TATACCCCTTAACACAGATATACCCCTCAAGGGGCGCTGTGCCACACATGGACCT 1566
Oy 355 -----ArgatgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsn 369
Db 1565 CACCCCCACAGCTAGAGAGGGGCGGTGCTG-----CTTGAG 1527
Oy 370 LysGluGluMetGluPhe----- 375
Db 1526 CTGCTGAGGTAGTAAGTAGGCTGCTTCTTCTGTGTGTCCTTCCCTCCAGTTCAT 1467
Oy 376 -----GlnProAsnGlnGluaspGluGluaspGluAlaHisAlaGlyIle 389
Db 1466 TGAGCTGGGGGTGGGGGTGGGACAGGACCTCAGAGCCACTGAGCATGAGAGGCTACT 1407
Oy 390 IlleglYArpHeuGluGluGlnSerHisAspHisHisProProArgAlaAsnSerArg 409
Db 1406 GAAGGGAACAACAAGAGGAGGAGACAGACAGAGGGAATGACAGGCCACCTGGAGA 1347
Oy 410 ThrLysLeuLeuTrp-----ProLysArgGluSerLeuLeu 421
Db 1346 AGAAGGTAAAGCTGGGGCTCCTCAAAAGTACAGGGGAGCCAAAGAGGGAAGAGCTCA 1287
Oy 422 His---GluGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
Db 1286 CACTGAGAGCGGTGCTCCC-----AGTGCAGGTGAG-----AGGGCCCATGAG 1245
Oy 441 AspAsnLysAlaLysTrpLysLysAlaValaspAlaPheLysSerGlyProLeu----- 458
Db 1244 GCGAGCGGGGTGGTCTGCTCCAG-----TACCCACCTTCCCTGCTTCA 1197
Oy 459 -----TygGlnArpProGlyTyrtYrSerAlaProGlnThrProLeu--- 472
Db 1196 CAAGACAGAGGGGTATCAAGACGAGGAGGTACAGGTCCACAGCTCACACCTTAGGC 1137
Oy 473 -----SerProThrProMetPhe 478
Db 1136 TGTGCTCACTCAGGCTACCTGCCCAACTTAGAGAGGTGAGGAGCTCAAAAGTTT 1077
Oy 479 PheProLeuGluPro-----SerAlaProSerLysLeuHisSerValThrGlyIleAsp 496
Db 1076 TTGATGATGCCGCCACTCGGTGAGGCCCCCTTCAAAAGTGCATCTGTTACTTTCAAATG 1017
Oy 497 ThrLysAspLysSerLeuLysThrValSerSerGly 508
Db 1016 AAAAAGAATTCAAGTGTGATGTCCAGGAAGAGGGG 981

```

RESULT 13

```

PCT-US94-12936-1/c
Sequence 1. Application PC/TUS9412936
GENERAL INFORMATION:
APPLICANT: The Johns Hopkins University
TITLE OF INVENTION: Tumor Suppressor WAF1
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW suite 1100
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12936
FILING DATE: 10-NOV-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A., Hoshel, Dale H.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.47878
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL LINE: GM
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 6p21.2
FEATURE:
NAME/KEY: CDS
LOCATION: 76..568
OTHER INFORMATION: /gene="WAF1"
PCT-US94-12936-1

```

Alignment Scores:

```

Pred. No.: 0.53 Length: 2121
Score: 98.50 Matches: 71
Percent Similarity: 29.81% Conservative: 22
Best Local Similarity: 22.76% Mismatches: 112
Query Match: 3.16% Indels: 107
DB: Gaps: 15

```

US-09-622-964-3 (1-585) x PCT-US94-12936-1 (1-2121)

```

Oy 263 AlaTyrtProGlyHisGluLeuValaspLeuValaspLeuPheThrPheLeuGlnPhePhe 282
Db 1793 GCAATGCTGCTGCTCAAGCTC----- 1773
Oy 283 PheTyrtValGlyTrpLeuLysValaspGluGlnLeu-----Asn 296
Db 1772 ---TACTGCGACCAATCTTAAATGTCTGACTCTTGTTCCTGCTATCAAAAGTGCAT 1716
Oy 297 PheGluGlyLysAspAspAspPheGluThrAsnTrpIleValaspArgAsnLeuGln 316
Db 1715 GAAGTGGGAGGAGGTGGGATGAGGAGGAGGAGTGTGCTGAGCATTT---TGAGGGCCAGT 1659
Oy 317 ValserleuLeuAlaValaspGluMetHisGlnaspLeuProArgMetGluProAspMet 336
Db 1658 GTCTCCCTC-----CTAGAAAGATCTACTACCCCATCA 1626
Oy 337 TytTPrasnLysProGluProGluProTyrtThrAlaLaserAlaGlnPhe----- 354
Db 1625 TATACCCCTTAACACAGATATACCCCTCAAGGGGCGCTGTGCCACACATGGACCT 1566
Oy 355 -----ArgatgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsn 369
Db 1565 CACCCCCACAGCTAGAGAGGGGCGGTGCTG-----CTTGAG 1527
Oy 370 LysGluGluMetGluPhe----- 375
Db 1526 CTGCTGAGGTAGTAAGTAGGCTGCTTCTTGTGTGTCCTTCCCTCCAGTTCAT 1467
Oy 376 -----GlnProAsnGlnGluaspGluGluaspGluAlaHisAlaGlyIle 389
Db 1466 TGAGCTGGGGGTGGGGGTGGGACAGGACCTCAGAGCCACTGAGCATGAGAGGCTACT 1407

```

```

QY 390 IleglyarPheleuglyleuGlnSerHisAspHisProProAlaAlaAsnSerArg 409
DB 1406 GAAGGGAAAGCAAGAGGGGAGGACACAGCAGAGGAAATTGACAGCCAGCTGAGAGA 1347
QY 410 ThrlyleuLeuTrp-----ProlysaArglySerLeuLeu 421
DB 1346 AGAAGGTTAGCTGGGGCTCTCTCAAAAGGTACAGGGAGCAAGAGGAAAGGCTCAA 1287
QY 422 His-----GluGlyleuProlysaAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
DB 1286 CACTGAGACGGCTCCCTCC-----AGTGCAGGCTCAG-----AGGGCCATGAG 1245
QY 441 AspAsnLysAlaTrpLysleuLysAlaValAlaPheLysSerGlyProLeu----- 458
DB 1244 GCGAGCGGGGGTGTCTGCTCCAG-----TACCCACCTTCCCTCCCTGCTTCA 1197
QY 459 -----TyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeu--- 472
DB 1196 CAAGACAGAGGGGGGTATCAAGAGCCAGGAGGTACAGGTCCAGCTCAGCCCTAGGC 1137
QY 473 -----SerProThrProMetPhe 478
DB 1136 TGTGCTCACTTACAGGCTACCTGCCCAACTTACAGAGGTGAGGGAGCTCCAAAGTTT 1077
QY 479 PheProLeuGluPro-----SerAlaProSerLysLeuHisSerValThrGlyLeuAsp 496
DB 1076 TTGATGATGCCCCCAGCTCGGTGAGGGCCCTTCAAGGTGCCATCTGTTACTTCTCAATG 1017
QY 497 ThrlyAspLysSerleuLysThrValSerSergly 508
DB 1016 AAAAAGAAATTCAGTCTGTAGTGTCCAGAAAGGGG 981

RESULT 14
PCT-US96-11886A-1/c
Sequence 1, Application PC/TUS9611886A
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND PRODUCTS FOR SELECTIVELY
TITLE OF INVENTION: KILLING OR INHIBITING THE GROWTH OF CELLS EXPRESSING THE
TITLE OF INVENTION: MAP1 GENE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: UNITED STATES OF AMERICA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/11886A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/001,248
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: W0461/7028MO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2121 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

```

? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: HOMO SAPIENS
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 76..570
PCT-US96-11886A-1

Alignment Scores:
Pred. No.: 0.53
Score: 98.50
Percent Similarity: 29.81%
Best Local Similarity: 22.76%
Query Match: 3.16%
DB: 5 Gaps: 15

US-09-622-964-3 (1-585) x PCT-US96-11886A-1 (1-2121)

QY 263 AlaTyrProGlyHisGluLeuAspLeuValProValPheThrPheLeuGlnPhePhe 282
DB 1793 GCATGCCCTGCTCCATAGCCTC-----
QY 283 PheTyrValGlyTrpLeuLysValAlaGluGlnLeuLe-----Asn 296
DB 1772 ---TACTGCCACCATCTTAAATGTCTGACTCTCTTGTCCGTGCTAATCAAGTGCAT 1716
QY 297 PropheGlyGlnAspAspAspPheGluThrAsnTrpLeuValAlaAspArgAsnLeuGln 316
DB 1715 GAATGGGGGAGGAGGGGGGTGATGAGGAAGGCTCGAGACAT---TGAGGGCCACT 1659
QY 317 ValSerLeuLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMet 336
DB 1658 GTCTCCTC-----CTAGAAGATCTACTCCCTCCATCA 1626
QY 337 TyrTrpAsnLysProGluProGlnProProTyrThrAlaAlaSerAlaGlnPhe----- 354
DB 1625 TATACCCCTAACACAGAGATATACCCCACTCAAGGGGCTGTGCCACACATGGACCT 1566
QY 355 -----ArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsn 369
DB 1565 CACCCCAACAGCTAGAGGAGGGGGCGGTGCTG-----CTTAG 1527
QY 370 LysGluGluMetGluPhe----- 375
DB 1526 CTGCTTAGTAGTAAGACTAGGCTGCCCTTCTTGTGTGCTCCCTCCCTCCAGTCCAT 1467
QY 376 -----GlnProAsnGlnLysAspGluGlnAspAlaHisAlaGlyLe 389
DB 1466 TGAGCTGGGGGTGGGGGTGGGACAGCCCTCAAGAGCCAGCTGAGAGGTACT 1407
QY 390 IleglyarPheleuglyleuGlnSerHisAspHisProProAlaAlaAsnSerArg 409
DB 1406 GAAGGGAAAGCAAGAGGGGAGGACACAGCAGAGGAAATTGACAGCCAGCTGAGAGA 1347
QY 410 ThrlyleuLeuTrp-----ProlysaArglySerLeuLeu 421
DB 1346 AGAAGGTTAGCTGGGGCTCTCTCAAAAGGTACAGGGAGCAAGAGGAAAGGCTCAA 1287
QY 422 His-----GluGlyleuProlysaAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
DB 1286 CACTGAGACGGCTCCCTCC-----AGTGCAGGCTCAG-----AGGGCCATGAG 1245
QY 441 AspAsnLysAlaTrpLysleuLysAlaValAlaPheLysSerGlyProLeu----- 458
DB 1244 GCGAGCGGGGGTGTCTGCTCCAG-----TACCCACCTTCCCTCCCTGCTTCA 1197
QY 459 -----TyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeu--- 472
DB 1196 CAAGACAGAGGGGGGTATCAAGAGCCAGGAGGTACAGGTCCAGCTCAGCTCAGCTTAGGC 1137

```



```

Db 1690 GAG-----GTAGAGAGACTCTATAGAGTCCCAAGGACATTACATGGAGAGACT 1743
QY 485 AlaProSerLysLeuHis-----SerValThr 493
Db 1744 GACCACACAGACTCAACCAACAGAGATATGAGATCAGATGCTCCAGCATTTACT 1803
QY 494 GlyIle-----AspThrLysAspLysSerLeuLysThrValSerSergly 508
Db 1804 TCTTCTTGATCAGACCTGAAAGAACGACGAGAGTCGTCCTTAATTGAGAAATTCACTA 1863
QY 509 AlaLysLysSerPheGluLeu-----LeuSergLysSerAspGlyAlaLeuMetGlu 525
Db 1864 GGAGATAGCAGACGAGGTCCAGAGACATGGTTAACACCAACAGCTCAGCAGCGG 1923
QY 526 -----HisProGluValSerGlnValArgArgLysThrValGluPheAsnLeuThrAsp 543
Db 1924 AGACGACAGACCGGCTTCAAAACCAAGTCACTGCTATGATTCAGAAATTAAATATCAGAAATT 1983
QY 544 MetProGluLysProGluAsnHisLeuLysGlu 554
Db 1984 GTTCACAGAGACACTGAAGACTCAACACGCGAA 2016

```

Search completed: July 27, 2003, 14:51:03
 Job time : 130.241 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:06:55 ; Search time 351.894 Seconds

(without alignments)
3743.798 Million cell updates/sec

Title: US-09-622-964-3

Perfect score: 3120

Sequence: 1 MTTTYSQVNAARLGSFSL.....TLKDHPYALENDPDAHS 565

Scoring table:

BLOSUM62	
Xgapop 10.0 , Ygapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delpop 6.0 , Delpext 7.0	

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

-MODEL=frame_p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09622966/rnat_22072003.101118.26006/app_query.fasta.1.2069
-BB=N.Geneseq.101002 -QFMT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO_MAP -LARGEQUERY -NEG.SCORE=0 -NAT -DSPLOCK=100 -LONGLOC
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database :

```

N.Geneseq.101002.*
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Query Match	Length	DB	ID	Description	
1	3116	99.9	2229	20	AA221227	Human CGICE short	
2	3047.5	97.7	2429	20	AA221228	Human CGICE long f	
3	1899	60.9	1916	20	AA221229	Mouse CGICE CDNA s	
4	1655.5	53.1	1857	22	ABAI4558	Human nervous syst	
5	1652.5	53.0	16650	22	ABAI4559	Human nervous syst	
6	1652.5	53.0	18530	22	ABAI4557	Human nervous syst	
7	1615	51.8	16125	20	AA221226	Human CGICE genom	
8	1571	50.4	1263	20	AAV99722	Human adult retina	
9	1265.5	40.6	7108	22	ABAI4556	Human nervous syst	
10	1045	33.5	1717	24	ABL89697	Human polynucleoti	
11	1009	32.3	2861	23	ABL10793	Drosophila melanog	
12	886	28.4	1608	23	ABL12609	Drosophila melanog	
13	851.5	27.3	10760	23	ABL10792	Drosophila melanog	
14	828	26.5	3592	23	ABL12636	Drosophila melanog	
15	801	25.7	3785	23	ABL12608	Drosophila melanog	
16	770	24.7	1526	23	AA572796	Drosophila melanog	
17	752	24.1	1345	23	ABL12637	Drosophila melanog	
18	694.5	22.3	2196	23	ABL16195	Drosophila melanog	
19	662	21.2	1350	22	AA576848	Human secreted pro	
20	523	16.8	305	22	ABAI1094	Human nervous syst	
21	509.5	16.3	7505	23	ABL16194	Drosophila melanog	
22	415	13.3	424	20	AAV86976	EST clone BK112.	
23	343	11.0	3592	23	ABL12636	Drosophila melanog	
24	324.5	10.4	2104	23	AA572795	Drosophila melanog	
25	264	8.5	392	22	AA11210	Human breast cance	
26	264	8.5	401	22	AA120109	Human breast cance	
27	257	8.2	527	22	AA117589	Human breast cance	
28	238	7.6	507	22	AA39532	Human breast cance	
29	238	7.6	507	22	AAK88431	Human breast cance	
30	233	7.5	3785	23	ABL12638	Drosophila melanog	
31	220	7.1	1662	23	AA570598	Drosophila melanog	
32	156.5	5.0	482	23	AA572794	Drosophila melanog	
33	119.5	3.8	65792	22	AA572854	DNA encoding novel	
34	118.5	3.8	2048	23	ABL04613	DNA encoding novel	
35	114.5	3.7	1630	21	AA575706	Genomic fragment #	
36	114	3.7	3822	23	ABL07431	Drosophila melanog	
37	113	3.6	20555	24	ABAI0443	Human prostate can	
38	112.5	3.6	9207	22	AA546790	Drosophila melanog	
39	112.5	3.6	9207	24	ABK28434	Streptococcus ther	
40	112	3.6	3606	23	ABL09625	DNA transcription	
41	111.5	3.6	2334	22	AA541571	Drosophila melanog	
42	111.5	3.6	2629	23	AA573197	DNA encoding nove	
43	111.5	3.6	2781	22	AA541032	DNA encoding nove	
44	111.5	3.6	3280	24	ABK83706	Human CDNA differe	
45	111.5	3.6	3426	19	AAV73473	Human USP CDNA #3.	

ALIGNMENTS

RESULT 1

AA221227 standard; CDNA; 2229 BP.

AA221227:

22-NOV-1999 (first entry)

Human CGICE short form CDNA sequence.

CGICE: Best's macular dystrophy; mutation: diagnosis; detection;

BMD: age-related macular dystrophy; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 105..1862
FT /tag= a


```
OY 501 SerLeuYsThrValSerSerGlyAlaIlyssSerPheGluLeuSerGluSerAsp 520
DB 1605 AGCTTAAAGACTGTGATGTTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCAT 1664
OY 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgGlyThrValGluPheAsn 540
DB 1665 GGGGCGCTTGATGGAGACCCCAAGATATCTCAATGAGAGAGAAACTGTGAGTTTAAAC 1724
OY 541 LeuThrAspMetProGluIleProGluAsnHisLeuYsGluProLeuGluGlnSerPro 560
DB 1725 CTGACGAGATATGCGAGATCCCGGAAATCATCCTCAAGAAACCTTGGACATTCACCA 1784
OY 561 ThrAsnIleHisThrThrLeuYsAspHisMetAspProGlyTrpPalaLeuGluAsnArg 580
DB 1785 ACCAAATATACACATCACTCACTCAAAAGATCATGATGATGCGCTTGGAAACAGG 1844
OY 581 AspGluAlaHisSer 585
DB 1845 GATGAGACCATTCCTCC 1859

RESULT 2
AAZ21228
ID AAZ21228 standard; cDNA; 2429 BP.
AC AAZ21228;
XX 22-NOV-1999 (first entry)
DE Human CGICE long form cDNA sequence.
XX CGICE; Best's macular dystrophy; mutation; diagnosis; detection;
KW BMD; age-related macular dystrophy; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 105..1412
FT /tag- a
FT /product- "CGICE short form protein"
PN MO9943695-A1.
XX 02-SEP-1999.
XX 22-FEB-1999; 99WO-US03790.
XX PF 25-FEB-1998; 98US-0075941.
XX PR 18-DEC-1998; 98US-0112926.
XX PA (MERI) MERCK & CO INC.
XX PA (UYOP-) UNIV UPPSALA.
XX PI Petukhin K, Caskey CT, Metzker M, Wadellius C;
XX DR MPI: 1999-540560/45.
XX DR P-PSDB; AAT29954.
XX PT Human and mouse polynucleotides encoding CGICE polypeptides
XX PS Claim 2; Fig 4; 67pp; English.
CC The present sequence represents the human CGICE cDNA sequence, which
CC when mutated is responsible for Best's macular dystrophy (BMD).
CC Polynucleotides encoding CGICE are useful for diagnosing whether a
CC patient carries a mutation in the CGICE gene. Normal and mutated
CC CGICE proteins are useful for identifying activators and/or inhibitors
CC of these proteins. In order to treat BMD, the CGICE gene offers a
CC simpler and cheaper method of diagnosing BMD without the need for the
CC presence of the patient. The gene may also be useful to discovering
CC the genetic cause of age-related macular dystrophy.
XX Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 4.28e-295 Length: 2429
Score: 3047.50 Matches: 583
Percent Similarity: 89.28% Conservative: 0
Best Local Similarity: 89.28% Mismatches: 2
Query Match: 97.68% Indels: 69
DB: 20 Gaps: 1

US-09-622-964-3 (1-585) x AAZ21228 (1-2429)
OY 1 MetThrIleThrTrpThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 105 ATGACCATCACTTACACAGCAAGCGCTAATGCCGCTTAGGCTCTTCCCGCTG 164
OY 21 LeuLeuYsTrpArgGlySerIleTrpYsLeuLeuTrpGlyIlePheLeu 40
DB 165 CTGCTGTGCTGGGGGCGGACCATCTTCAACCTGTATATGCGAGTCTTAATCTTCTG 224
OY 41 LeuYsTrpTrpIleIleArgPheIleTrpArgLeuAlaLeuThrGluGlnLeu 60
DB 225 CTGCGTACTACATCATCCGCTTATATAGCTGGCCCTCACGGAAACACACACTG 284
OY 61 MetPheGluYsLeuThrLeuTrpYsAspSerTrpIleGlnLeuIleProIleSerPhe 80
DB 285 ATGTTGAGAAACTGACTGTATGCGACAGCTACATCCAGCTCATCCCATTCCTTC 344
OY 81 ValLeuGlyPheTrpValThrLeuValValThrArgTrpPheAsnGlnTrpGluAsnLeu 100
DB 345 GTCTGGGCTTCTACAGTGACGTGTGTACCCGCTGGTGAACCATGACGAACTG 404
OY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyYsAspGluGln 120
DB 405 CCGTGGCCGACCGCTCATAGACCTGTGTGGGTCTGCTGCAAGCAAGCAAGCA 464
OY 121 SerArgLeuLeuArgTrpThrIleArgTrpAlaAsnLeuGlyAsnValLeuLeu 140
DB 465 GCGCGGCTGCTGGGGGCGGACGCTACCGCTACGCAACCTGGGCAACGTCATCTC 524
OY 141 ArgSerValSerThrAlaValTrpYsArgPheProSerAlaGlnHisLeuValGlnAla 160
DB 525 CGCAGCGTCAGACCCGCACTTACAGCCGCTTCCACGCGCCACGACCTGGTCAAGCA 584
OY 161 GlyPheMetThrProAlaGluHisYsGlnLeuGluYsLeuSerLeuProHisAsnMet 180
DB 585 GCGTTTATGACTCCGCGCAGAACACAGCAGTTGGAGAACTGAGCTTACACACACATG 644
OY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetYsAlaTrpLeuGlyArg 200
DB 645 TTCTGGGTCCCTGGGTGTGTGGTTCGCAACCTGTCAATGAAAGCGCTGGAGGTGCA 704
OY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgTrpGln 220
DB 705 ATCCGGAGACCCATCTCTGCTCCAGACCTGCTGCAAGACAGATACACCTTGGTACTCG 764
OY 221 CysGlyHisLeuTrpAlaTrpAspTrpIleSerIleProLeuValTrpThrGlnValAl 240
DB 765 TGTGACACCTGTATAGCTTACGACTGATTAATCCCATGCTGTGTATACAGAGTGCTG 824
OY 241 ThrValAlaValTrpSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB 825 ACTGTGGGGGTGTACAGCTTCTCTGACTGTCTGTAGTTGGGCGGAGTTCTGAACCA 884
OY 261 AlaYsAlaTrpProGlyHisGluLeuAspLeuValAlaProValPheThrPheLeuGln 280
DB 885 GCCAAGGCTTACCCCTGCGCAGATGAGCTGACCTGTGGCCCGCTTACAGTCTCTGCGAG 944
OY 281 PhePhePheTrpValGlyTrpLeuYsVal----- 290
DB 945 TTCTTCTTCTATGTGTGGCTGTGAAGCT-GGGCCTTCCAGGCGCTGTGGCTGGAG 1003
OY 290 ----- 290
DB 1004 GCATGCCAGAGGGGTGATGGCCACAGCTGCTTGTAGACGAGATGACGTGTCAAGAAAG 1063
```

QY	290	-----	290
Db	1064	GAAGGTCTACGGGTAGAAAGCAGCGAGCGTGGTGGCCACACACTGTAAATCCAGCTAC	1122
QY	291	-----	291
Db	1124	TCGGGACCGCTGAGGACGAGAGAAATCCGTTGAAACCCGGGAGCGGAGGTTGTGGTGGCAGAG	1183
QY	293	GlnLeuIleasnProPheGlyIuaAspAspAspPheGluThrAsnTrpIleValAsp	312
Db	1184	CAGCTCATCAACCCCTTTGGAGAGGATGATGATGATTTTGGACCAACTGATTTGCGAC	1243
QY	313	ArgAsnIleuAlaSerLeuIleuAlaValAspGluMetHisGlnAspLeuProArgMet	332
Db	1244	AGGAATTTGACAGTGTCCCTGTTGGTGTGGATGAGATGACACAGAGCCGTGCTTGAGTG	1303
QY	333	GluProAspMetIlyTrpAsnLysProGluProGlnProProTyrThrAlaAlaSerAla	352
Db	1304	GACCCGGACACTGTACTGGAATAAAGCCCGAGCCACAGACCCCTTACAGAGCTGCTTCCCC	1363
QY	353	GlnPheArgArgAlaSerPheMetGlySerTrpPheAsnIleSerLeuAsnLysGluIu	372
Db	1364	CAGTTTCCGTGACAGCTCTTTATGGCTGTCACACTTCAACATCAGCTCAACAAAGAGAG	1423
QY	373	MetGluPheGlnProAsnGlnIuaSpGlnIuaSpAlaHisAlaGlyTleGlyArg	392
Db	1424	ATGGAGTTCCACAGCCCAATCAGAGAGCGAGAGATGCTCAACGCTGGCATATTGGCCGC	1483
QY	393	PheLeuGlyLeuGlnSerHisAspHisIshProAlaGlnAlaAsnSerArgThrLysLeu	412
Db	1484	TTCTTAGGCTCAGAGTCCCATGATCACCATCTCTCCAGGGGAAACTCAAGAGCAACCACTA	1543
QY	413	LeuTrpProLysArgIuSerLeuIleuHisGlnGlyLeuProLysAsnHisLysAlaAla	432
Db	1544	CTGTGGCCCAAGAGGAATCCCTCTCCACAGAGGCCCTGCCCAAAACACAGAGCGACC	1603
QY	433	LysGlnAsnValArgGlyGlnIuaSpAsnLysAlaTrpLysLeuLysAlaValAspAla	452
Db	1604	AAACGAAACGTTAAGGGCCAGGAAACAAACAAGGCCGTGAAGCTTAAAGCTGTGAGCGCC	1663
QY	453	PheLysSerGlyProLeuTyrGlnArtrProGlyTyrTyrSerAlaProGlnThrProLeu	472
Db	1664	TTCAAGTCTGGGCCACTGTATCAGAGGCCAGGCTACTAGTGGCCCAACAGAGCGCCCTC	1723
QY	473	SerTrpThrProMetPhePheProLeuGluProSerAlaProSerLysLeuHisSerVal	492
Db	1724	ACCCCACTCCCATGTTTCTCCCTTAAACATCAGAGCGCCGTCAAAAGCTTCAACAGTGTG	1783
QY	493	ThrGlyIleAspThrLysAspLysSerLeuIlysthrValSerSerGlyAlaLysLysSer	512
Db	1784	ACAGGCAATAGACACCAAAAGCAAAAGCTTAAGAAGCTGTGAGTCTGGGCGCAAAAGAT	1843
QY	513	PheGluIleuLeuSerGluSerLysArgGlyAlaLeuMetLuhisProGluValSerGlnVal	532
Db	1844	TTTGAAATGCTCTCAGAGAGGAAAGGGGCCCTGTATGAGACACCAGAAAGATCTCAAGTG	1903
QY	533	ArgArgLysThrValGluPheAsnLeuThrAspMetProGluIleProGluIuaAsnHisLeu	552
Db	1904	AGGAGAAACACTGTGAGGTTTAACCTACGAGATTTGCCAAGATCCCGGAAATATCACTTC	1963
QY	553	LysGluProLeuGluGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAsp	572
Db	1964	AAAGAACTTTGGAAACATACCCAAACCAATACACTACACTCAAAAGTACATGAT	2023
QY	573	ProTyrTrpAlaLeuGlnIuaAspArgSerGluAlaHisSer	585
Db	2024	CCTTATTGGGCTTGGAAGAAACAGGAGTAAAGCCACATTC	2062
RESULT 3			
AA621229			
AA621229	standard; cDNA; 1916 BP.		
XX			

```

AC  AA21229;
XX
DT  22-NOV-1999 (first entry)
XX
DE  Mouse CG1CE cDNA sequence:
XX
KM  CG1CE: Best's macular dystrophy; mutation; diagnosis; detection;
XX  BMD; age-related macular dystrophy; ss.
XX
OS  Mus sp.
XX
FH  Location/Qualifiers
FT  11..1666
FT  CDS /tag=a
XX  /product= "CG1CE protein"
XX
PN  W09943695-A1.
XX
XX  02-SEP-1999.
XX
XX  22-FEB-1999; 99MO-US03790.
XX
XX  25-FEB-1998; 98US-0075941.
XX  18-DEC-1998; 98US-0112926.
XX
XX  (MER1 ) MERCK & CO INC.
XX  (UYUP-) UNIV UPSALA.
XX
PI  Petrukhin K, Caskey CT, Metzker M, Wadelius C;
XX
DR  WPI, 1999-540560/45.
XX  P-PSDB; AAY29955.
XX
XX
XX  Human and mouse polynucleotides encoding CG1CE polypeptides
XX
XX  Claim 2; Fig 8; 67pp: English.
XX
XX
XX  The present sequence represents the mouse CG1CE cDNA sequence, which
XX  when mutated is responsible for Best's macular dystrophy (BMD).
XX  Polynucleotides encoding CG1CE are useful for diagnosing whether a
XX  patient carries a mutation in the CG1CE gene. Normal and mutated
XX  CG1CE proteins are useful for identifying activators and/or inhibitors
XX  of these proteins, in order to treat BMD. The CG1CE gene offers a
XX  simpler and cheaper method of diagnosing BMD without the need for the
XX  presence of the patient. The gene may also be useful to discovering
XX  the genetic cause of age-related macular dystrophy.
XX
SQ  Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 other;

Alignment Scores:
Pred. No.: 4.29e-180 Length: 1916
Score: 1899.00 Matches: 378
Percent Similarity: 74.32% Conservative: 59
Best Local Similarity: 64.29% Mismatches: 101
Query Match: 60.87% Indels: 50
DB: 20 Gaps: 10

US-09-622-964-3 (1-585) x AA221229 (1-1916)

QY 1 MetThrIlePheThyThrHisSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
   |||||
Db 11 ATGACTATACCTACCAACAACAAGATGACCAATGGCCGCGTCTGCTTCCTGCTCC 70
   |||||

QY 21 LeuLeuGlyStrpArgGlySerIleTyrlsLeuLeuTyrlGlyGluPheLeuIlePheLeu 40
   |||||
Db 71 CTCTGTGCTGGCGGAGGACGATCAACAAGCGTGTGTGAGAAATTCCTTGTTCATA 100
   |||||

QY 41 LeuGlyStrpArgGlySerIlePheArgPheIleTyrlArgGluAlaLeuGlnGluGlnGlnLeu 60
   |||||
Db 131 TTCCTCTACATATTCATCCGCTGGACTCTACAGAAATGGCTTCTCTGAGTATGACGACGCTG 190
   |||||

QY 61 MetPheGluIleuLeuTyrlsTyrlsAspSerTyrlIleGlnLeuIleProIleSerPhe 80
   |||||

```

```

Db      191 TTGTTGAGAGAGTGGCTGTACTGGACAGCTACATTCACTCATCCCTAATATCCTTC 250
QY      81 ValLeuGlyPheTyrValIThrLeuValValThrArgTrrPaspnGlnTyrGluAsnLeu 100
      251 GTTCTGGGTTTCTATGTATTGATGTTGAGCCCGGTGGTGGAGCCGATGACGAACTTG 310
QY      101 ProTrrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAspGluGln 120
      311 CGGTGGCCGACCCGCTCATGATCCAGGTGCTGAGTTCCTGCGTGGAGGGCAAGATGAGAA 370
QY      121 SerArgLeuLeuArgTrrThrLeuIleArgTrrAlaAsnLeuGlyAsnValLeuIleu 140
      371 GGCGTTTGTGCTGGGGCAGCGCTCATCCGCTAGCCATCCGCGCCAAAGTCTCATCTCG 430
QY      141 ArgSerValSerThrAlaValTyrTyrAspPheProSerAlaGlnHisLeuValGlnAla 160
      431 CGGAGCATGACGACCTGGGTGTACAAAGCCCTTCCACCTTCACACCGCTGGGTAGCA 490
QY      161 GlyPheMetThrProAlaGlnHisLeuGlnLeuGlyLeuSerLeuProHisAsnMet 180
      491 GGTATTATGACCCATGGGACATAGACATGCTGCAGAAAGTTGGGCTACACACACACACA 550
QY      181 PheTrrValProTrrValIThrPheAlaAsnLeuSerMetTyrAlaTrrLeuGlyArgT 200
      551 TTCTGGGTGCTCCGTGGGTGTGTGTCACACTTGTCAATGAAAGCCATCTTGGAGGTGCA 610
QY      201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgTrrGln 220
      611 ATCCGGGACACCGCTGCTGCTCCAGAGCCGTGATGAAAGAGGTGTGACTTGGCTACTCG 670
QY      221 CysGlnHisLeuTyrAlaTyrAspTrrIleSerIleProLeuValTyrThrGlnValAl 240
      671 TGTGACAGACTGTATGCTTACAGACTGGATTAATCCCATTTGTGTACACAGAGGTGGG 730
QY      241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
      731 ACAGTGGCATGTATACGCTTTTCTCTTCATGCTGATGGGAGGACAGTTCTGTGACCCA 790
QY      261 AlaTyrAlaTyrProGlnHisGluLeuAspLeuValIleProValPheThrPheLeuGln 280
      791 AACAAAGACTACCCAGGCGCATGAAATGATGCTGGTGTGCTGTCTTACAAATCCTCGAA 850
QY      281 PhePhePheTyrValGlyTrrPheLeuValAlaGluGlnLeuIleAsnProPheGlyGlu 300
      851 TTCTATTCTCATAGGGCTGGCTGAGGTGGCAGAAAGCTCATCAACCTTTGGGGGAG 910
QY      301 AspAspAspAspPheGluThrAsnTrrIleValAlaAspArgAsnLeuGlnValSerLeuLeu 320
      911 GACGATGATGATTTTGAAGACTAACTGATGATGACAAACSTGCAGGTGCTCCCTGTG 970
QY      321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProArgMetTyrTrrAsnLys 340
      971 TCCGTGATGGGATGACACAGAACTTCCCTCCATGGAACSTGACATGTAATGAAACAG 1030
QY      341 ProGluProGlnProArgTrrTrrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
      1031 GCAGCGGCTCAGCGCGCTCAGACAGTCTTCTCCAGGTGCTGCCGCGCATTCCTTCATG 1090
QY      361 GlySerThrPheAsnIleSerLeuAsnLysGluIleMetLurPheGlnProAsnGlnGlu 380
      1091 GGCTCCACCTTCACATCAGCCTTAAAGAAAGAAAGATTAAGCTTGGTCAAAAAGAGAG 1150
QY      381 -----AspGluGluAspAlaGlnAlaGlyIleLeuGlyArgPheLeuGlyLeu 396
      1151 GCTGACACGAGTAAAGAAAGAGTGGGTATAGCAGCAGCAGCTAGAGCTCTTAAAGAGCTG 1210
QY      397 GluSerHisAspHisThrProArgAlaAsnSerArgTrrHisLeuLeuTrrProLys 416
      1211 CAACCCAAAACATACATCTTCCCTGAAAGACTTAAAGCAAAACATATGTGTGTAAAG 1270
QY      417 ArgGluSerLeuHisGluGlyLeuProLysAspHisLysValAlaLysGlnAsnVal 436
      1271 AACCCCTCTCTC-----GAAAGCCAGGTAAAGAT-----GCCAAACAG 1309

```

```

QY      437 ArgGlyGlnGluAspAsnLysAlaTrrPrrLeuValAlaValAspAlaPheLysSerGly 456
      1310 AAAAACAGAAAGAT-----GTCTGAAATTTAAAGGCTGTGACATCTTGAATGTGT 1363
QY      457 ProLeuTrrGlnArgProGlyTrrTyrSerAlaProGlnThrProLeuSerProThrPro 476
      1364 CCAAGGTTTAAAGAGAGAGGCTCCCTTGTGGCCACAGACCCAGCAGCAGC----- 1414
QY      477 MetPhePheProLeuGluProSerAlaProSerTrrLysLeuHisSerValThrGlyLea 496
      1415 -----CACCTTACTAGCAGATGACAGCAGCCTCCAGT-----TCAGACAGAGT 1456
QY      497 ThrLysAspLysSerLeuTrrValSerSerGlyAlaLysLysSerPheGluLeu 516
      1456 ----- 1456
QY      517 SerGluSerAspGlyAlaLeuMetGlnHisProGluValSerGlnValArgArgLysThr 536
      1457 -----GATGGGCGCTCCACAGATTTACAAAGAAATGTGTACATGAAAGAAACT 1507
QY      537 ValGluPheAsnLeuThrAspMetProGluIleProGluAsnHisLeuLysGlu---Pro 555
      1508 GTGGAGTTTAACTTG---AACATTCAGAGAGCCCCACAGACATCTTCAACAGCGCGCT 1564
QY      556 LeuGluGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAspProTrrTrr 575
      1565 TTGGACCAAGNTGTCAACCAATATACAGCCTGTAAATGAAGACATGACAGAGTCTAT--- 1621
QY      576 AlaLeuGluAsnArgAspGluAla 583
      1622 -----CCCTACAGGATGAAGCT 1639
Db      1622 -----CCCTACAGGATGAAGCT 1639

```

RESULT 4
ABAI4558 standard; DNA; 18537 BP.

ABAI4558;
23-JAN-2002 (first entry)

Human nervous system related polynucleotide SEQ ID NO 6889.
Human; nocturnal; neuroprotective; cytosolic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery; antiparkinsonian; antischling; antianemic; antitarrtic; cancer; antipneumatic; hepatocytic; cerebroprotective; antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal; antiparasitic; cardiac; immune disorder; cardiovascular disorder; neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
Homo sapiens.
W0200159063-A2.
16-AUG-2001.
17-JAN-2001; 2001MO-US01334.
31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180528.
PR 24-FEB-2000; 2000US-0184564.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225275.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0225681.
 PR 22-AUG-2000; 2000US-0225681.
 PR 22-AUG-2000; 2000US-0225686.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232339.
 PR 14-SEP-2000; 2000US-0232339.
 PR 14-SEP-2000; 2000US-0232339.
 PR 14-SEP-2000; 2000US-0232339.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.

PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249222.
 PR 17-NOV-2000; 2000US-0249223.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249226.
 PR 17-NOV-2000; 2000US-0249227.
 PR 17-NOV-2000; 2000US-0249228.
 PR 17-NOV-2000; 2000US-0249229.
 PR 17-NOV-2000; 2000US-0249230.
 PR 01-DEC-2000; 2000US-0250300.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-541565/60.
 Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 useful for preventing, diagnosing and/or treating nervous system
 cancers and metastases -
 Disclosure; SEQ ID NO 6889; 1701ip + Sequence Listing; English.

The invention relates to novel genes (ABAI1004-ABA21534) and proteins
 (ABAI4678-ABBI8001) useful for preventing, treating or ameliorating
 medical conditions e.g. by protein or gene therapy. The genes are
 isolated from a range of human tissues disclosed in the specification.
 The nucleic acids, proteins, antibodies and (ant)agonists are useful
 in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 and ovarian cancer and other cancers of the adrenal gland, bone, bone
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX

SQ Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 other;

Alignment Scores:

Prod. No.: 3,31e-154 Length: 18537
 Score: 1655.50 Matches: 529
 Percent Similarity: 20.29% Conservative: 0
 Best Local Similarity: 20.29% Mismatches: 5
 Query Match: 53,068 Indels: 2073
 DB: 22 Gaps: 7

US-09-622-964-3 (1-585) x ABA14558 (1-18537)

```

OY 51 ATGLeuAlaLeuThrGluGlnGlnLeuMetPheGluLysLeuThrLeuTyrCysAsp 70
DB 4743 AGCGTGGCCCTCAGCGAAGAACACAGCTGATGTTGAGAACTACTCTGATTGCGAC 4802
OY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
DB 4803 AGCTACATCCAGCTCATCCCATTTCTTCTGCTGCTGAGTCCCTCTGCTGCTGT 4862
OY 82 ----- 82
DB 4863 CCGGGTCCCTGTCGCCGCCAGGCTCCAGACAGCCAGGGAGATCAAGAGACTGCG 4922
OY 82 ----- 82
DB 4923 GCAAGGGCTGGGGAGGGGGGGGGGAAAGCCAGCGAGTCCGCCCTCTGTAGAG 4982
OY 82 ----- 82
DB 4983 AAGAGTGGGACTGACGACGACAGAGAAAGTAAGATTAGAGTTCCTGCGCT 5042
OY 82 ----- 82
DB 5043 TAGCAATGAAAAACCCCATTTTCTGAGGGAAGCGCTGACATCATGTCCCTGAGCCCTG 5102
OY 82 ----- 82
DB 5103 CGCGGAGGGGAGGGGGTCTGGCGGATTTCTGGGACGACAGGGGAGCCCGGGTGACA 5162
OY 82 ----- 82
DB 5163 GAACCCCTGGGGCTTCGCGCTTCATCGAGGCTCTGCGCTCGCTCCCGAGCGC 5222
OY 82 ----- 82
DB 5223 CTTCAGAGAGGCTGGGGGCTAGGCGCGCTCGCAGAGAAAGCTGAGAGCGAGGCA 5282
OY 82 ----- 82
DB 5283 CGCGGGGCGCTGGGCCCTGGGCTCTGGCCGACAGCTGGCCCTCGCCCTCGCCCGCC 5342
OY 83 -----GlyPheTyrValThrLeuValValThrArgTyrTrpPheAsnIleTyrGlu 98
DB 5343 CCCTCTCTGCCAGGCTTCTACGTGACGCTGCTGCTGAGCCCGCTGGGGAACAGTACGAG 5402
OY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
DB 5403 AACCTCCCGTGGCCCAACCGCTCATGAGCTGTGTCTGGGCTTCGTGAGGCAAGAGAC 5462
OY 119 GluGlnSerArgLeuLeuArgArgTrpLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
DB 5463 GAGCAGAGCGCGGTGTGGCGGCGACGTCATCCGCTACGCCCACTGGGCAACGTGCTC 5522
OY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal 158

```

```

DB 5523 ATCTTGGGAGCGCTCAGCAGCAGGCTCTACAGCGCTTCCCGAGCGCCGACCTGTG 5582
OY 159 Gln----- 159
DB 5583 CAAGCAGTGTGGCGGACCGGAGAACAGGGGAGACACCGGCGACAGCCAGGGCCAGAT 5642
OY 159 ----- 159
DB 5643 GGGCGCGCAGAAATGAAATGGTGTGAGCAAGTCCCGGACTCGGGGATTTGGGT 5702
OY 159 ----- 159
DB 5703 GGAGCCAGAGATGGGGTGTGTCAAGATTTTGGGGGTCCAATTGGCGGAGACAGTCTGG 5762
OY 159 ----- 159
DB 5763 TGTCTGAAGTGGGGCGAGGCCAGGACCCACCTCCGAGATGAGGAGTCTGAGGCGAGG 5822
OY 159 ----- 159
DB 5823 CTAAGGACCTTGAAGGATATGAAAGAGGGTACGCGTTGGAACTGTGAGTACT 5882
OY 159 ----- 159
DB 5883 AGGCTACTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5942
OY 159 ----- 159
DB 5943 GGACATTGTCTGTGACACCCCTCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6002
OY 159 ----- 159
DB 6003 ACCAGGCTAGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6062
OY 159 ----- 159
DB 6063 GAATGCGCTTGAAGCCAGCTGTTGAGAGCCCGCTGAGCAATGAGGAGACCCCATC 6122
OY 159 ----- 159
DB 6123 TCTCAAAAAACATTTAAAAATTAGCAGGCGATGTGGGTGTGCTGTAGTCTGAGCTGA 6182
OY 159 ----- 159
DB 6183 GATACGGAGGCTGAGGCGAGGAGATCATTGAGCCAGCAATTCCAGCTGCACTGCC 6242
OY 159 ----- 159
DB 6243 TAAGATCGACCGCTGCACTCAACCTCGGTGACAGACAGACACCTTCTGTGAAATA 6302
OY 159 ----- 159
DB 6303 AATTAATACCTCTCCCATATGCTCAGCCAGAAACAGACCTAGTAGTCTCAGAAATT 6362
OY 159 ----- 159
DB 6363 TTTTGTGTGAAGAAAGAGATGGCAAGAGAGTGTGCTGATAGTCTATAGTCAAGAG 6422
OY 160 -----AlaG1 161
DB 6423 TGGCGGCAATCCCTTGTGAGGTTCTCCACCCAGCCCTTCTTCACTCCACTCTCAGAG 6482
OY 161 yPheMetTrpProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetPh 181
DB 6483 CTTTATGACTCCGGCGCAACAGACAGAGTGGAGAACTGAGCTTACACACAACTGTT 6542
OY 181 eTrpValProTyrValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyLysArg 201
DB 6543 CTGGGTGCCCTGGGTGTGTGTTGCAACCTGTCAATGAAGAGGGTGTGAGAGTGCAT 6602
OY 201 eArgAspProIleLeuGlnSerLeuAsn----- 212

```

Db	6603	CCGGACCCCTATCTCTCTCCAGAGCCTGCTGAACGTGAACCCCACTGTACAGACAGGGCTG	6662
Oy	212	-----	212
Db	6663	CCGACAGATGGGAAGGGTTGTGGTCCACAGAAACAAGTTTCTTACAAAGAAAGCCTT	6722
Oy	212	-----	212
Db	6723	GGGCCCTGAGGGTCTTCGAGAGCCGAGGTGGGGTTGCAGATCTTTTCCAAACAGCA	6782
Oy	212	-----	212
Db	6783	TCCACAGCCGAGGTGTCTCCCTTCTCAGAGCCCTCCCTCTTCAAGTGTGAGGT	6842
Oy	212	-----	212
Db	6843	CCGGTTCCCTTTTGTATAGATGAGGAAGCTGAGACAAAGAAGTTTATGAGCTTCCCA	6902
Oy	212	-----	212
Db	6903	TGGCCACACAGCCAGGAAATGACCATATGATACCAGGCCCTGTGTACTGGAGAAAGTGG	6962
Oy	212	-----	212
Db	6963	GGGCCAGCCACAGGTGGGGGACGTGTGTTCAGAAACCCATCCCTCTTCTGCCCCC	7022
Oy	213	--GluMetAsnThrLeuValArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIleSerI	232
Db	7023	AGCAGATGAACACCTTGGGTACTCGTGTGACACACTGTATATGCTTACACTGATTTGTA	7082
Oy	232	LeProLeuValTyrThr-----	237
Db	7083	TCCCACTGCTGTATACACAGGTGAGACTAGGCTGTGTGAGCTGCCTTTTGGGAAACTG	7142
Oy	237	-----	237
Db	7143	AGGCTAGAAAGCAAGAACAGCTGGGGTGGGAAGGGCTCACCTAGAGCTAAGTGCC	7202
Oy	237	-----	237
Db	7203	TCCCTGGGAGTGGGTCTCACACTTTGAAGTTGGGTCTGGACTTTGAAGTCCAAAGTTCT	7262
Oy	237	-----	237
Db	7263	AAGAGTCCAGGCTCTCTGCTGGCCCAAGTCCAGTAGAGGCAATGTGATTATCCCATTTTA	7322
Oy	237	-----	237
Db	7323	AAGAGAGGTTGGCGGGGACAGTGGCTCATCTGTAATCCACACTTTGGGAAGCTGA	7382
Oy	237	-----	237
Db	7383	GGCAGGTGATCACCTGAGTCAAGAGTTCGAGACCAAGCCTGGGCCAATGGTGAACCC	7442
Oy	237	-----	237
Db	7443	CATCTTACTGAAATAACAGATTAGCTGTGTGTGTGCACGCTGTATATCCAGCTAC	7502
Oy	237	-----	237
Db	7503	TTGGGAGGCTGAGGCAGAGAAATCGCTTGAACCCGGGAGGTGAGTTGCATGAGCTGA	7562
Oy	237	-----	237
Db	7563	GATCATGCACACTGCTCCAGCCTGGGCGACACAGCAAGACTGTGTCTCAACAAACAAA	7622
Oy	237	-----	237
Db	7623	CAAAACAACAACAACAACAACAAGGGTTAACAGAGCCCTTAAGTACATTAAGTG	7682
Oy	237	-----	237
Db	7683	TGCAGTCACAAAGGCCCTTGGTCTTCCTGTCTTCACAGTCCAGCCCTTGGAGCATCTGG	7742

Oy	238	-----G	238
Db	7743	ATTTCAGGGTTCCCACTAGACCCTTTCCTACACATCCTCTCTCTCCCTCC	7801
Oy	238	InValValThValAlaValItyrSerPhePheLeuthrcYstLeuValIslYArGslInPhel	258
Db	7803	AGGTGGTAGCTGTGGCGGTGTACAGCTTTCTCTACTTGCTGTAGTGGGGCGGAGTTTC	7867
Oy	258	euaNPrOAlalysalatyProglYnlsGluleuAspleuValProvalPherhp	278
Db	7863	TGAACCCACCAAGGCCCTACCCCTGGCCATGACGTGACCTGTGCCCCGTCTACAGT	7922
Oy	278	helenslnPhephePheTyValIslYTrleu	288
Db	7923	TCCTCAGTTCCTTCTTAATGTGGCTGAGTAAAGTGGGCCTCCACAGGCCCTGTGG	7983
Oy	288	-----	288
Db	7983	GCTGAGCATGCGCCAGAGGGGTCAATGGCCACAGCTGTGAGACAGAGATGACGTGC	8044
Oy	288	-----	288
Db	8043	AGAAGAGGAGGCTCACGGGTAGAAAGCAGCCAGCGTGTGGCCACACCTGTAAATCC	8102
Oy	288	-----	288
Db	8103	CAGTACTCGGGAGGCTAGAGCGAGAGAAATGCTTGAACCCGGGAGCGGAGTTGTGT	8167
Oy	288	-----	288
Db	8163	GAGTTAGATCGTGCCACTGACCTCAGCTGGGCCAAAAGATGAACCTATCTCANAA	8222
Oy	288	-----	288
Db	8223	ACAACACACACAAAAAAGCCCTTAAGTTCAGAAGCCCCTTGCCCTTAGAAGCGAG	8282
Oy	288	-----	288
Db	8283	AGCGGACACACCTCTCTTATTTCAAGATGCTGTGGGCTGTCTTGTTCCTCACCTCAG	8342
Oy	288	-----	288
Db	8343	TGGCTTGCCAGGTAATTCCTCCACACAGCAATACTCGAAGAGATGTTCTGAATCA	8402
Oy	288	-----	288
Db	8403	CACAGTTTTCTCTCCACCTCTTATCTTCTCTCTCTGTGGCCACACCTCTCTCTC	8462
Oy	288	-----	288
Db	8463	CCCTTCATCCTCTTATTTTGGTAATGGGGGTGAAGTCTGTCTCTGCCCTTCC	8522
Oy	288	-----	288
Db	8523	TGTCACTGTGACACACACACACACACACACACACACACACACATTCCTAT	8582
Oy	288	-----	288
Db	8583	TCTCTAAATTTCCCTGTGCCCCCAGTATCTTGTGTTCTGTGAGATCAAAMCAATCA	8642
Oy	288	-----	288
Db	8643	CACITTTATGCTGAATAATCTCCAGGGTGCCCCAGTGCCGTGCAAGATGTCCCTGGACC	8702
Oy	288	-----	288
Db	8703	CCTAAGCGACAGCGGTGTCAACCTCTGGGGGCTTTGTTAGGGCATTTTAGAGTTGTCTAT	8762
Oy	288	-----	288
Db	8763	CCAGGAATCTGCCACCTAGACTGCCCTTATGTTACGCCACGCTTCAGTATATATCTCG	8822


```
QY 288 ----- 288
Db 8823 TTGCATGAATGAATAAATTATGCAACTCCAGTAGATACATGAGGTGAGATAAAGCA 8882
QY 288 ----- 288
Db 8883 GTGACTCAGCCGAGTGATACACTCAGGGCAGACCTGTGGTGTTCAGGAGAGACTGGCTC 8942
QY 288 ----- 288
Db 8943 AGAAGATTAGAGGGCTGTGTCCAGAAAGTGTGTGGTCCCAAGTGTGGGGGGCTGG 9002
QY 288 ----- 288
Db 9003 AGCCCTAAACTCTGCTTTGAAGACAGTGTGCAGGAGAGAGGCTTCATGGGTGTGGA 9062
QY 288 ----- 288
Db 9063 AATAGCAGCAGCTGAGTTTAAAGGGGGAAGCTGCTTTGAGGAGTTCTGCTGAGGGTT 9122
QY 289 ----- 289
Db 9123 TACAGAGCTCAGCTGTCCCAAGGTGGCAGAGAGCTCATCAACCCCTTTGGAGAGGAT 9182
QY 302 AspAspAspPheGluThrAsnTrpIleValAspArgAsnLeu ----- 315
Db 9183 GATGATGATTTTGAGACCAACTGATTTGTCCAGAGGAATTTGACAGTATGGGAGAGGGA 9242
QY 315 ----- 315
Db 9243 GAGAAACCAATACATGAGACTTCCCAAAAGTGAGACCAAAAGAGAGACCCCATGTTCTG 9302
QY 315 ----- 315
Db 9303 TAGGAGAGCTCAGACGTGAATGATCAACCTTCCCTCCTCCTGACAGCAGTCAAT 9362
QY 315 ----- 315
Db 9363 CACTCAGAGATTCTCAGCTCAATCTTTGAGGCTGCAGGAGGACCAACCATCTCCCAATT 9422
QY 315 ----- 315
Db 9423 CACAGCAGGAAACTGAGGTCAGAGAGAGGAGAGATCTTCCAAGTCATCAGGACA 9482
QY 316 ----- 316
Db 9483 TACAGGTCTGCTGGATGATCTTCTGTGGACTTCTTGTCCCTGGTGACCAAGT 9542
QY 317 LserLeuLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTy 337
Db 9543 GTCCCTGTGGCTGTGATGATGATGACACGAGACCTGCTCGATGAGACCGGACATGTA 9602
QY 337 rTPAsnLysProGluProGlnProProtyrThrAlaAlaSerAlaGlnPheArgArgAl 357
Db 9603 CTGGATTAAGCCCGAGCCACAGCCCTCAGACAGCTGCTCCGCCAGTTCCGTGAGC 9662
QY 357 aserPheMetGlySerThrPheAsnIle ----- 366
Db 9663 CTCCCTTATGGCTCACCCTTCAACATCAGGTGTGCGCCAGAGCCAGGGGGCTGGGTGGA 9722
QY 366 ----- 366
Db 9723 AGCCCTCCTAGTCAGAGGGGTCTGCCTAGAACTAGAAATAGCACTAGTTAATGCAATACA 9782
QY 366 ----- 366
Db 9783 GGTGCTTCAGTAAGTGTACAGCAGCTGTACTATGCTCTTTATTAACATAACTATTATTTT 9842
QY 366 ----- 366
Db 9843 TCCTCCCAATATTCGTGGTTGTATCCCAAGTTTTCAGATATTAAGTACAGGTTGAG 9902
QY 366 ----- 366
QY 288 ----- 288
Db 9903 AGAGATGAATGTTCCAAAGGCCACATAGCTACCAATGATGCTACTCGAAGGAC 9962
QY 366 ----- 366
Db 9963 AGCCTATGATCAGTATGATCAGTGGAAAGCTTGAAGACCTGGCTTGTCAATCCAGAACTATG 10022
QY 366 ----- 366
Db 10023 TTTTCTTTTCTTTTGTAGACAGTATCTGCTCTGTGCGCCAGTTGAGCGCAGTGGCT 10082
QY 366 ----- 366
Db 10083 GATCTGCTACTGCAACCTCCGCTCTCTGGTTCAAGTATTCCTGCTTGAAGCTC 10142
QY 366 ----- 366
Db 10143 CCCAGTAGCTGGATTTACAGGTGCCCAACACCAACTGCTAATTTTGTACTTTTACT 10202
QY 366 ----- 366
Db 10203 AGAGATAGGTTTACAGCATGTGGCCAGGCTGTGTCTCCAACTCCTGACAGATACTGCC 10262
QY 366 ----- 366
Db 10263 CGCTTGGCTCCCAAAATGCTGGAATTAATAGTGTCAAAACTATGTTTCTGTAAAGCT 10322
QY 366 ----- 366
Db 10323 ACGATGCTTGATGGAGAGTGAAGTGGGTTCCCTGGATGGGGGAGGGGACCAAGT 10382
QY 366 ----- 366
Db 10383 CCCAGCAGGACGACGAGCCATCAAGTACTCTCTGAATGACTTGTCTTACGAGTAA 10442
QY 366 ----- 366
Db 10443 AGGGCTCAGGACCAACCAAGCAGCAGCAGACTATTCACCATGTCTCCACTTCCCTATTG 10502
QY 366 ----- 366
Db 10503 CATCTAATTCCTCTTGTAGCTGCATGGGCTGAAGGCTATCCAGCTGTCTTCTCTC 10562
QY 366 ----- 366
Db 10563 CCAGCACAACAGAGTTGAAGTGCCTTGGAGAGTGTGGGACATGTCAAGGTTCACTACT 10622
QY 366 ----- 366
Db 10623 CAAGGTTTCTTCAAGGTATCCAGTGTGTCTGCTTCTTTTCTTTTCTTTTCTTTT 10682
QY 366 ----- 366
Db 10683 TAAAGGAGTTCACTCTTGTGCCAGAGCTGAGTGAGTGCAATATCTCGGCTCAC 10742
QY 366 ----- 366
Db 10743 TCCAACTCCGCTCCAGATTCAGCAATTCCTGCTCAGCTCCGTAAGTACGTGG 10802
QY 366 ----- 366
Db 10803 AATTAGTCCAGCAGCAGCAGCCCGGCTAATTTTGTATTTTGTAGAGACACTTCA 10862
QY 366 ----- 366
Db 10863 CCATTTGGCAGAGGTGTCTCGAAGTCTGACTCAGGTATTCACCCCTCTAGCCTC 10922
QY 366 ----- 366
Db 10923 CCAAGTGTGGATTAATATGTGTAGCAGCTGTGCTGTGTCTTCTTTTAAAGAC 10982
QY 366 ----- 366
```

Db 10993 CAATATCTCTACTAGACTGCATCGATTTAACAGCTCTATAGATAGTGTGAGGANG 11042
 QY 366 ----- 366
 Db 11043 GTTGGGAAGGTGATCAATTAAGAGCTGAGGCTTCTTAGTCAAGAAATTCTTGAGG 11102
 QY 366 ----- 366
 Db 11103 ATGACTTTGAGCCCTACATGGTCTGTACCCAGCAGCTGAAGTTGTGAGGGGTGGGA 11162
 QY 366 ----- 366
 Db 11163 GGGCTGAACAGAACGATTAAGCATAGACCTTGTCTCAAGAAATGCACAAATTTATGGA 11222
 QY 366 ----- 366
 Db 11223 GGGAGCTCAAAACCAAGCTCAAACTCTGGATACAAAGTACAAAGTACTGATGTCCAGA 11282
 QY 366 ----- 366
 Db 11283 AAGGAGACAGAACATGNAACAGTCACTTTGTCTGCTGGAGGGGCTTCCAGCTGG 11342
 QY 366 ----- 366
 Db 11343 GTCGTGAGCTGAGCCATGGAACATGGAGAAATCTGAACCTTGGCAAGGCGACCAATAC 11402
 QY 366 ----- 366
 Db 11403 TCTCTGTAGATAGCTTTCCCTTGCAGGGTAAAGTCTGGGGCTCCCGGGATGCTGTT 11462
 QY 366 ----- 366
 Db 11463 GCTAGAGTCAATTTCTCTTGTGTGATGTCACTCCAGTGGACACAAATCTCTGG 11522
 QY 366 ----- 366
 Db 11523 CATTTGCCCAAGTCACTCATGGGCTCATCTGAACCATCATGCCAGGCGACCAAGTCTT 11582
 QY 366 ----- 366
 Db 11583 CTGACTGCTGAGTGAAGGGGTTTACAGGGGAAGTGAATGATGAGAGGCTTTACAG 11642
 QY 366 ----- 366
 Db 11643 CCAGCGGGGGTGGTTCGGGGGTTGGATTAACTCTGTGTCAAGAGGAATCAACAACA 11702
 QY 366 ----- 366
 Db 11703 GTGAGTGAAGCTGGGCTGGAGGATCACCGGAGGTACAGCAGATCAGAGAGAGGT 11762
 QY 366 ----- 366
 Db 11763 GAGAGCTGGGGCATGTGTAGGAAGAGGTGTGGCTTGGGCCAATGAGAGAG 11822
 QY 366 ----- 366
 Db 11823 GAGCGGGGTAAGGGAAGTAAGGCCAGTGTGCTCTTGTCCAGCTCAGCCCT 11882
 QY 367 ----- SerLeuAnLysGluGluPheGlnProAsnGlnI 380
 Db 11883 GCATCTCTCTGTTCTTTCACAGCTGACAAAGAGAGATGAGCTTCCAGCCCAATCAGGA 11942
 QY 380 uAspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAs 400
 Db 11943 GGAACGAGGAGATGCTCAAGCTGAGCATCTTGGCCGCTCCAGCCCTCAGTCCACAGA 12002
 QY 400 pHisHisProPArgAlaAsnSerArgThrIleuLeuTProGlySerGluSerIle 420
 Db 12003 TCACCATCTCTCCAGGCGCAACTCAAGGACCAAACTACTGTCGCCCAAGAGGAATCCCT 12062
 QY 420 uLeuHisGluGlyLeuProLysAsnHisLysAlaIleLysGlnAsnValArgGlyGlnI 440
 Db 12063 TCTCCACAGGCGCTGCCCCAAACACACAGCAGCAACAGAAAGCTTAAAGGCGCAGGA 12122

QY 440 uAspAsnLysAlaIleArgLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrG 460
 Db 12123 AGACAAAGAGGCTCGGAAGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGTATCA 12182
 QY 460 nArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPrometPhePhePr 480
 Db 12183 GAGGCGAGGCTACTACAGTGGCCACAGACACCCCTCAGGCCCACTCCCATGTTCTTCCC 12242
 QY 480 OLeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLy 500
 Db 12243 CCTGAAACCATCAGCGCCGCTCAAGCTTCAAGTGTCAAGGCTAGACACCAAGACAA 12302
 QY 500 sSerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAs 520
 Db 12303 AAGCTTAAGACTGTGATGCTTGGGGCCCAAGAAAGTTTGATTTGCTCTCAGAGACGA 12362
 QY 520 pGlyAlaLeuMetClnHisProGluValSerGlnValArgArgLysThrValGluPheAs 540
 Db 12363 TGGGCGCTGATGAGACCCACAGAAATGTCTCAAGTGAAGAGGAAACTGTGGAGTTAA 12422
 QY 540 nLeuThrAspMetProGluIleProGluAsnHisLeuLysGluProLeuGluGlnSerPr 560
 Db 12423 CTTGAGCGATATGCCAGAGATCCCGGAAATTCACCTCAAGAAACCTTTGGAACAATCAC 12482
 QY 560 oThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTyrPAlaLeuGluAsnAr 580
 Db 12483 AACCAACATACACACTCACTCAAAAGATCAGATGATCTTATTTGGGCGCTTGAAGAAACG 12542
 QY 580 gAspGluAlaHis 584
 Db 12543 GTCTGTCTCCAC 12555
 Db 12555
 RESULT 5
 ABA14559
 ID ABA14559 standard; DNA; 16650 BP.
 AC ABA14559;
 XX
 DT 23-JUN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide seq ID NO 6890.
 XX
 OS Homo sapiens.
 PN WO200159063-A2.
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214866.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.

CC collitis (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 other;

Alignment Scores:

```

Pred. No.:      5,63e-154      Length:      16650
Score:          1652.50        Matches:      529
Percent Similarity: 20.31%     Conservative: 0
Best Local Similarity: 20.31%  Mismatches:  5
Query Match:    52.96%        Indels:      2071
DB:             22            Gaps:         7

```

US-09-622-964-3 (1-585) x ABA14559 (1-16650)

QY	51	ATGLeuAlaLeuThg1ug1ug1nleuMetPheg1u1ySLeuThleuThyCTyasp	70
Db	4743	AGCGTGGCCCTCAGGAGAAACAACAGCTGATGTTTGAAACTGACTCTGATTTGCGAC	48020
QY	71	SerTyrlleGlnLeuIleProIleSerPheValLeu	82
Db	4803	ACCTACATCCACACTCATCCCATTTCTCTCGTGGGTGAGTTCCCTCTGTGGCTGTT	48652
QY	82	-----	82
Db	4863	CCGGGTCCCTGTGGCCGCCCAAGCTCCAGACAGCCAGGGAGGATCAAGAGAGCTGCG	49322
QY	82	-----	82
Db	4923	GCAMGGGCTGGGAGAGGGGGGGAACGCCAGCGAGGTCCGGCCCTCTCTGTAGGG	49822
QY	82	-----	82
Db	4983	AAAGTGGGACTGCAGCCAGAGAAACTGAATTAGACTTAAAGTAAAGACTCTGGCGT	5042
QY	82	-----	82
Db	5043	TAGCAATGAAAACCCCATTTTCTGAGGGAAGCGCTGACATCATGTCTCCGTGAGCCCTG	5102
QY	82	-----	82
Db	5103	CGCGGAGAGGAGGGGCTGTGGCGAATTCTGGGACCAAGAGGGGACCCCGGGTGACA	5162
QY	82	-----	82
Db	5163	GAACCTTTGGGGCTCTCGCGCTTCATCGAGGCTCTGCCTCTCGCTCCCGACGCG	5222
QY	82	-----	82
Db	5223	CTTCAGAGAGGGCTAGGGGGCTAGGCCCGCTTCGACAGCAAAAGCTGAGAGCCGAGCAT	5282
QY	82	-----	82
Db	5283	CGCCGGGCGTGGGCCCTGTGGGCTGTGGCCGAGCGTGGCCCTCGGCCCTCGGCCGCG	5342
QY	83	-----GlyPheTyrlValThleuValThaArgTrpPabaGlnTyrglu	98
Db	5343	CCCTCCCTGCCCAAGGCTTCTACGAGACCTGTGTGACCCTGTGGTGAACACAGTACGAG	5402
QY	99	AsnLeuProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAsp	118
Db	5403	AACCTGCGGTGGCCGACCCCTCATGAGCTGTGTGTGGGGCTTCTCGAAGGCAAGGAC	5462
QY	119	GluGlnSerArgLeuLeuArgTrpLeuIleArgTyrlaAsnLeuGlyAsnValIleu	138
Db	5463	GAGCAAGCGCGCTGCGCGGCGCACAGCTCATCCGCTAGCGCCAACTGGGCAACGCGTCTC	5522
QY	139	IleLeuArgSerValSerThrAlaValTyrlValArgPheProSerAlaGlnHisLeuVal	158

Db	5523	ATCTCGACGAGGTCAGCACCGCAGTCTACAAAGCGCTTCCCGAGCGCCACACTG	5582
QY	159	Gln	159
Db	5583	CACGACGGTGGCGGACCGGGAGCAACGGGGAGGACCGGGCAGAGCCAGGGCCGAGAT	5642
QY	159		159
Db	5643	GGGGCGGCGAGGAATGGAAAGATGGGTGGAGCCAAAGTCCCGGACTCGGGGATTGGT	5702
QY	159		159
Db	5703	GGAGCGGAGAGTGGGGTGTGTGTCAGATTGGGGGTCCATTGGCGGGACAGAGTGGG	5762
QY	159		159
Db	5763	TGTCGAAAGTGGGGCGAGGCCAGAGCCACCTCCGAGTAGAGTCTGAAGCAGG	5822
QY	159		159
Db	5823	CTAAGACCCCTTGAGGATTAATGAAAGAGAGGTGACGGCTTGGGAACCTGTGAGGTACT	5882
QY	159		159
Db	5883	AGGCTACTTCCCTGCTGCGCTTGCCTTGTATCTCGGTTTCCACTCTGGAGTATG	5942
QY	159		159
Db	5943	GGACATTGGTCTGACACCCCCCTCAGCTGGCTGACCTGTGCTTGTAAATAGACAG	6002
QY	159		159
Db	6003	ACCCAGGCTAGCGGTGGTGGCTCTCGCGCTGTAAATCCAGTGTTAAGAGCAAGGTGG	6062
QY	159		159
Db	6063	GAAAGTCGCTTGAGCCCAAGCTGTTTGAGACGCCCTTGAGACACATAGCAGACCCCATC	6122
QY	159		159
Db	6123	TCTACAAAACATTAAAAATTAGCAGGGCATGGTGGCGTGTGCCCTGTAGTCTGAGGCTGA	6182
QY	159		159
Db	6183	GTAATGGGAGGCTGAGCGCAGAGGATCATTTAGCCGACAGTGTCCAGGCTGCAATGCGC	6242
QY	159		159
Db	6243	TAAAGTGCACACCGGTGACTCCAACTCGGTGACAGACGACAGACCTTCTCGAATAA	6302
QY	159		159
Db	6303	AATAAATACCTGCCACATGTCAGCCGAGAACAGCACATAGTAGTGTCTGAAAATT	6362
QY	159		159
Db	6363	TTTTGTGTGTAAGAAGAAAGATGGCAAAAGAGTGTGAGTTCCTAATAGTCAAGCAG	6422
QY	160		161
Db	6423	TGCGGACCATCCCTTCTGCAAGTTCGCCACCCAGCGCTTCTTCACTCCACTGCGAGG	6482
QY	161		181
Db	6483	CTTTATGACTCCGGCAGAACACACAGAGTTGGAGAAACAGACCTTACACACACATGTT	6542
QY	181	CTTATGACTCCGGCAGAACACACAGAGTTGGAGAAACAGACCTTACACACACATGTT	201
Db	6543	CTGGGTGACCTGGGTGTGGTTTGGCCAACTGTCAATGAAGCGGTGGCTTGGAGGTGCAAT	6602
QY	201	CTGGGTGACCTGGGTGTGGTTTGGCCAACTGTCAATGAAGCGGTGGCTTGGAGGTGCAAT	212

Db	6603	CCGGGACCCCTATCCCTGCTCCAGAGCCCTGCTGAACGTGAGCCCACTGTACAGACAGGGCTG	6662
Oy	212	-----	212
Db	6663	CCGCAAGTGGGNAAGGCTTGTGTCACACAGAAACAAGTTTCTTACAAAGAGAGCCCTT	6722
Oy	212	-----	212
Db	6723	GGGCCCCCTGAGGGTCTTCCGAGAGCCGAGGTGGGGTTGCAGAACTTTTCCAACAGCAA	6782
Oy	212	-----	212
Db	6783	TCCACAGCCCGGAGGTGGTCCCTTCTCAGAGCCCTCCCTCTTCCAAAGTGTGAGGT	6842
Oy	212	-----	212
Db	6843	CCTGTTCCCTTTTGATAGATGAGAGCTGAGACACAAGAGGTTTATGAGCTTCCCA	6902
Oy	212	-----	212
Db	6903	TGGCCACACAGCCAGGAATGACCATAGTATACAGGCCCTGGTACCTGGAGAGAGGTGG	6962
Oy	212	-----	212
Db	6963	GGGGAGCCCAAGGCTGGGGGACGGTGGTTCAGAACCCATCCCTCTCTGCCCCC	7022
Oy	213	--gIuMeLanThrLeuAlrGthrGlnCysGlyHisLeuTyralaTyrlsPrlIseLrI	232
Db	7023	AGGAGATGAAACACCTTGCGTACTCATGTGTGGACACCTGTATGCTACGACGATTAAGTA	7082
Oy	232	LeProlLeuValItyrThr-----	237
Db	7083	TCCCACTGGTGTATACACAGGTGAGACTAGGCTGTGAGGCTGCCCTTTGGAAACIG	7142
Oy	237	-----	237
Db	7143	AGGCTAGAGAGCACAAGAGACAGCTGGGGTGGGAAGGCTCACCTAGAGCTAAGTGAC	7202
Oy	237	-----	237
Db	7203	TCCCTGGAGATTGGTCCACACATTTGAAGTTGGGTCTGGACTTTGAAGTCCCAATTCT	7262
Oy	237	-----	237
Db	7263	AAGAGTCCAGGCTCTGCTGCGCCAGTCCAGTAGAGGCATGTGATTATCCCATATTA	7322
Oy	237	-----	237
Db	7323	AAGAGAGTTGGCCGGGACACAGTGGCTCATGCTGTATCCACACTTTGGAAAGCTGA	7382
Oy	237	-----	237
Db	7383	GGCAGGTGATCATCTGAGGTCAAGACTTGAGAACCAAGCCTGGCCAAACATGTGAAACC	7442
Oy	237	-----	237
Db	7443	CATCTCTACTGAAATAACAGAAATTAGCTGTGTGTGTGCACGCGCTGTATCCAGCTAC	7502
Oy	237	-----	237
Db	7503	TTGGGAGGCTGAGGCCAGGAAATCGCTTGAACCCGGGAGGTGAGATTGCAGTGAAGCTGA	7562
Oy	237	-----	237
Db	7563	GATCATGCACTGCACTCCAGCCTGGGCGCACACAGCAAGACTTGTCTCAACAAACAAA	7622
Oy	237	-----	237
Db	7623	CAACAAACAAACAAAGGGGTTAACAGAGCCCTTAAGTCAATTAAGTGTGCAAGTTC	7682
Oy	237	-----	237
Db	7683	AGAACAAAGCCCTTGCTCTCTGCTCTCAGACTCCCAAGCCCTGAGAGCACTGATTTACAG	7742

QY	238	-----GlnValVal1_240
Db	7743	GTTCCACCAGGCCCTTGTGCTACACATCCTCGTCCTCCCTCCTCCAGTG 7802
QY	241	ThrValAlaValtyrserPhepheLeuThrCysLeuValJglyargInPheLeuAsnPro 260
Db	7803	ACGTGGCGGTGCACAGCTTCCTCCGACTGCTGTAGTTGGGGCGCAATTCTGAACCCA 7862
QY	261	AlaLysAlaItyr-ProGlyAsnLeuAspLeuValValProValPheThrPheLeuGln 280
Db	7863	GCCAAAGGCTACCCCTGGCCATGAGTGGAACTGTGTGGCCCTTCACGTTCTGTGAC 7922
QY	281	PhephePheTyValJgLyTrpleu-----288
Db	7923	TTCCTCTCTAgttgctggcctgaagaggcgccctccaggcccctggcgctggag 7982
QY	288	-----288
Db	7983	CATGGCCAGAGGGGTGATGCGCAGCATGCGCTGAGACGAGATGCAGTGTCAAGAAAG 8042
QY	288	-----288
Db	8043	AAGTCTACGGGTAGAAAACAGCAGCGCGTGTGGCGCACACTGTATCCAGCTACT 8102
QY	288	-----288
Db	8103	CGGAGGCTGAGGCGCAGAGAATCGTGTGAACCCGGAGGCGGAGTTGTGTGACTTGAG 8162
QY	288	-----288
Db	8163	ATCTGCCACTGCACCTCCAGCGCTGGCAAAGAAATGAACCTATCTATAAAAAACAAC 8222
QY	288	-----288
Db	8223	AACAACAAAACAAAGCCCTAAGTTTCAGAAAGCCCCCTGCCCCTTAGAAGCAGAGCGGAC 8282
QY	288	-----288
Db	8283	CACCTCTCTATTATCAAGATGCTGTGGGCTGTCTGTCTCCTACCTCAAGTGGCTGT 8342
QY	288	-----288
Db	8343	CGAGGTATTCCTCCGCCAACACACAGCAATCTCCGAAAGATGTTCTGAATCACACAGTTT 8402
QY	288	-----288
Db	8403	TCATCCACCTTATCTTATCTTCTCTCTGTGTGGCCACCACACTCTCTCCCTCTTA 8462
QY	288	-----288
Db	8463	CCTTCTTATTTTTGGTAATGGGGGTGAAGTCTGTCTGTGCCCCCTTCCTGTCACTG 8522
QY	288	-----288
Db	8523	TGACTTCTTATCTCTCT 8582
QY	288	-----288
Db	8583	AAATTTCCCGCTGCCCGCCAGTTATCTTTGGTTTCTGCAGATCAAAAAANAATCACACTTT 8642
QY	288	-----288
Db	8643	TATGCTTGAATTTCTCCAGGCTGCCCACTGGCCTCAAGATGTCCCCCTGAGCCCTAAG 8702
QY	288	-----288
Db	8703	GCAGACGGGTGTCACCTCTTCGGGCTTTGTAGGCAATTTTAGAGGTGTATTCACAGA 8762
QY	288	-----288
Db	8763	ATCTGCCACCTAGACTGCCCTTTAGTTTTCAGCCCAAGCTTCAGTATATATCTGTTCAT 8822

Db 10983 TCCTACTAGACTGCATCGAGTTTAACACTAGCTATAGATAGTGTGAGGAGGGTTGGG 11042
 QY 366 -----
 Db 11043 AAGGTATCAAAATGAGGCTGAGAGCTTGCTTAGTGCAGAAACATTTCTGAGAGATGACT 11102
 QY 366 -----
 Db 11103 TTGAGCCCTACATGCTGTGATCCCGACAGCTGAGGTTGTGAGGGGTGGGAGGGCTG 11162
 QY 366 -----
 Db 11163 AAAACAGAACGATTAAGCATAGACCTTGCTGTCAGAGAAATGCACATTTATGAGAGGAGC 11222
 QY 366 -----
 Db 11223 TCNAACCAAGTCTCAAACTCTGAGATACAGGTACAAAGTACTGATGTCAGAAAAAGG 11282
 QY 366 -----
 Db 11283 ACAGAACATGAAACAGTATCTTGTGCTGGGAGGGGCTTCCAGCTGGGTCTG 11342
 QY 366 -----
 Db 11343 AGCTGAGCATGGAACATGGGAAGAATCTGAATTGGGCAAGGCGAGCCATCTCTG 11402
 QY 366 -----
 Db 11403 GTAGATTAAGCTTTCCCTGTCAGAGGTAAAGTCTGGGGCTCCCGGAGTCTGTTGTAAG 11462
 QY 366 -----
 Db 11463 AAGTCAATTTCTCTTTGTGATGTCACCTCCAGTTGGAACCAAAATCTGTCATTC 11522
 QY 366 -----
 Db 11523 CCAGATCATCTATGAGGCTCATCTGAACCACTATGCCAGGCGACAGTCTTGTGACT 11582
 QY 366 -----
 Db 11583 GCCCTGAGTAGGGGTTTACAGGGGAAGTAATGATGAGAGAGCCCTTACAGCCAGGC 11642
 QY 366 -----
 Db 11643 GGGGTGCTGGGGGGTGTGATGTTAACTGTGCTCAAGAGGAATCAACAAAGTGAAG 11702
 QY 366 -----
 Db 11703 TGAGCTGGGCTGAGAGGATCACCGGAGGTACAGAGACATGAGAGAGAGGTGAGGC 11762
 QY 366 -----
 Db 11763 TGGGCGATGGTGAAGAGAGGCTGTGCTGGCTTGGCCAAAGTGAAGAGAGAGCGG 11822
 QY 366 -----
 Db 11823 GGGTAAGGAGAGTAAGGCCAGGTGTTGGTCTTTGTCACACTGCGCTCAGCTCATCT 11882
 QY 367 -----
 Db 11883 CCTGTTCTTTCCAGCTGAAACAAAGAGAGATGAGTTCAGCCCAATCAAGAGAGAGCA 11942
 QY 382 UGUUASPAAHLSAAGLYLLEGLYARGPHEUGLYLEUGLINSERHISASPHLSH 402
 Db 11943 GGAGGATGCTACGCTGCGATCATGCGCGCTCTCTAGGCTTGACGTCCTCATGATCAACA 12002
 QY 402 SPRTORATGAAASERATGTHLYSLEULEUTRPROLYSARGLISERLEULEUH 422
 Db 12003 TCCTCCAGGGGCAAACTCAAGAGCAAACTACTGTGGCCAAAGGGGAATCCCTTTCCTCA 12062
 QY 422 SGLUGLYLEUPRPROLYSASPHLSLSALAALALYGLASALVALARGELYGLINGLUASPAS 442
 Db 12063 CGAGGGGCTTGCCCAAAAACCAAGGAGAGCCAAACAGAAACGTTAGGGGCCAGGAAGACAA 12122

QY 442 nLysalatrPLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrgInArPr 462
 Db 12123 CAAGGCTGTGAGGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGTATCAAGAGCC 12182
 QY 462 OGlyTyrrTySerAlaProGlnThrProLeuSerProThrPheMetPheProLeuGI 482
 Db 12183 AGGCTACACAGTGTGCCCAAGAGAGCCCTCAGGCCCACTCCATGTTCTTCCCTTAGA 12242
 QY 482 uProSerAlaProSerLysLeuHisSerValThiGlyLLeaspThrLysAspLysSerLe 502
 Db 12243 ACCATCACCGCCGCTCAAGGCTTCAAGTGTCAAGGATACAGCAACCAAGCAAAAGCTT 12302
 QY 502 ULysThrValSerSerGlyAlaLysLysSerPheGIuLeuLeuSerGIuSerAspGI 522
 Db 12303 AAAAGCTGAGTCTGCGGGCCAAAGAAAGTTTGAAATGCTCTCAAGAGAGCGTGGGC 12362
 QY 522 ALeuMetGIuHisProGIuValSerGIuValArgArgLysThrValGIuPheAsnLeuThr 542
 Db 12363 CTTGATGAGACACCCCAAGATATCTCAAGTGAAGAGAGAAACGTGAGCTTAACTGAC 12422
 QY 542 RAspMetProGIuLLeProGIuAsnHisLeuLysGIuProLeuGIuGlnSerProThrAs 562
 Db 12423 GGATATGCCAGAGATCCCCGAAATCACCTCAAGAAACCTTTGGAAACATCACCAACAA 12482
 QY 562 nLeHisThrThrLeuLysAspHisMetAspProTyrrTPalaLeuGIuAsnArgAspGI 582
 Db 12483 CATACACACTACACTCAAGATCATGATGATCTTATTTGGGCTTGAAAGAGGTCTGT 12542
 QY 582 uAlaHis 584
 Db 12543 CCTGCAC 12549
 RESULT 6
 ID ABA14557 standard; DNA; 18530 BP.
 AC ABA14557;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 6888.
 XX
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; vitruclide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antischizoid; antinaemic; antiarthritis; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PE 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225211.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227189.
 PR 30-AUG-2000; 2000US-0227909.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241825.
 PR 20-OCT-2000; 2000US-0242221.

PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249247.
 PR 17-NOV-2000; 2000US-0249255.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251859.
 PR 08-DEC-2000; 2000US-0251988.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI: 2001-541565/60.
 Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 useful for preventing, diagnosing and/or treating nervous system
 cancers and metastases -
 Disclosure: SEQ ID NO 6888; 1701pp + Sequence Listing; English.
 The invention relates to novel genes (ABAI1004-ABAI21534) and proteins
 (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
 medical conditions e.g. by protein or gene therapy. The genes are
 isolated from a range of human tissues disclosed in the specification.
 The nucleic acids, proteins, antibodies and (ant)agonists are useful
 in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 and ovarian cancer and other cancers of the adrenal gland, bone, bone
 marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX

SQ Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 other;

Alignment Scores:

Pred. No.:	6.61e-154	Length:	18530
Score:	1652.50	Matches:	529
Percent Similarity:	20.31%	Conservative:	0
Best Local Similarity:	20.31%	Mismatches:	5
Query Match:	52.96%	Indels:	2071
DB:	22	Gaps:	7

US-09-622-964-3 (1-585) x ABA14557 (1-18530)

```
OY 51 ATGLeuAlaLeuThrgLugInuLeuMetPheGluLysLeuThrLeuTyCysAsp 70
   |||||
Db 4743 AGGCTGCCCCCTCAGGAGAAACAAGCTGATGTTGAGAAACTGCTGTATTCGAC 4802
OY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu-----
   |||||
Db 4803 AGCTACATCAGACTCATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4862
OY 82 -----
   |||||
Db 4863 CCGGCTCCTGTCGCGCCGAGCTTCAGACAGCCAGGAGGAGATCAGAGAGCTGCG 4922
OY 82 -----
   |||||
Db 4923 GCAAGGGCTGGGAGGGGGGGGAGGAGCCAGCGCAGTGGCGCTCTGTAGAG 4982
OY 82 -----
   |||||
Db 4983 AAAGTGGGACTGCAGCAGCAGAAACTGAGTTAGAGTTAGAGTACGTCCTGCGGT 5042
OY 82 -----
   |||||
Db 5043 TAGCAATAAACCCCATTTTCTGAGGAGAGCGCTGACATCATGTCCTGAGCCCTG 5102
OY 82 -----
   |||||
Db 5103 CGCGGAGAGGAGGGGGCTGCGGATTTCTGGACACGACAGGGGAGCCCGGGGTGACA 5162
OY 82 -----
   |||||
Db 5163 GAACCTTGGGGCTGCGCGCTCCATGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5222
OY 82 -----
   |||||
Db 5223 CTTCAGAGAGGCTGGGGGCTAGCCCGCTGCGACAGAAAGCTGAGAGCCGAGCAT 5282
OY 82 -----
   |||||
Db 5283 CGCGGGCGCTGGCCCTGCGCGAGCTGCGCCCTGCGCCCTGCGCCCTGCGCCCTGCG 5342
OY 83 -----
   |||||
Db 5343 CCGCTCCCGCCAGGCTTCTAGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5402
OY 99 AsnLeuProTrrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAsp 118
   |||||
Db 5403 AACCTGCGCTGGCGCCGAGCCGCTCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5462
OY 119 GluGlnSerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
   |||||
Db 5463 GACCAAGGCGCGCTGCTGCGCGCCAGCTCATCCGCTAGCCCAACCTGGGCAACGTCGCT 5522
OY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnIleVal 158
```

```
Db 5523 ATCTGGCGACGCTCAGCACCAGCTCTACAGCGCTTCCCGAGCCGACCTGCTG 5582
OY 159 Gln-----
   |||||
Db 5583 CAAGCAGGTGGGGGAGACCGGGAGCAACGGGGAGGACCGGGAGCCAGGCGCGAGAT 5642
OY 159 -----
   |||||
Db 5643 GGGCGCGCAGGAATGAGATGGTGAGCCAAAGTCCCGGACTCGGGGATTTGGGT 5702
OY 159 -----
   |||||
Db 5703 GGAGCCAGAGTGGGGTGTGTGTCAGATTTGGGGGTCCAAATTGGCGGGAGACAGTGGG 5762
OY 159 -----
   |||||
Db 5763 TGTCTGAAGTGGGGGAGGAGCCAGAGCCACCTCCGAGAGTAGAGTCTGAGCGAGG 5822
OY 159 -----
   |||||
Db 5823 CTAAAGACCTTGAAGGATATGAAAGAGGGTACGGCTTGGAACTGTGAGGTACT 5882
OY 159 -----
   |||||
Db 5883 AGGCTTACTTCCCTCTGCGCTTGGCCCTTGTATCTCCGTTTCCACTCTGAGGTATG 5942
OY 159 -----
   |||||
Db 5943 GGACATGTGCTGTACACCCCTCAGCCTGAGCTGAGCTGCTGCTGCTGCTGCTGCTG 6002
OY 159 -----
   |||||
Db 6003 ACCGAGCTAGGCGTGGTGGCTCTGCGCTTGAATCCAGTCTTAGAGAGCAAGGTGG 6062
OY 159 -----
   |||||
Db 6063 GAAGTCGCTTGAAGCCAGCTGTTGAGACGCCCTGAGCAACATAGGAGAACCCCATC 6122
OY 159 -----
   |||||
Db 6123 TCTCAAAAACATTAAAAAATTAGCAGGAGCATGTGGGTGCTGAGTCTGAGGCTGA 6182
OY 159 -----
   |||||
Db 6183 GTATCGGAGGCTGAGGAGGAGGATCATTTGAGCCAGCAGTTCAGAGCTGCAGTGGCG 6242
OY 159 -----
   |||||
Db 6243 TAAGATGCACCGCTGCATCAACCTCGGTACAGAGAGCAGACCTTTCTGGAATA 6302
OY 159 -----
   |||||
Db 6303 AATTAATACCTGCCACATGCTCAGCCAGAAACAGCAGCTAGTAGGTCTGAGAAATT 6362
OY 159 -----
   |||||
Db 6363 TTTTGTGTGAAGAAAGAGATGGCAAGAGGTGTGATGTTCTATAGTCTGACAGG 6422
OY 160 -----
   |||||
Db 6423 TGGCGGCATCCCTTTCGAGGTTCTCCACCCAGCCGCTTCTTCACTCCACTCTCAGG 6482
OY 161 -----
   |||||
Db 6483 yPheMetThrProAlaGlnHisLysGlnLeuGlyLysLeuSerLeuProHisAsnMet 181
   |||||
Db 6483 CTATTAGACTCCGGGAGAAACAGAGAGTTGGAGAACTGAGCTTACACACAACTGTT 6542
OY 181 eTrrPValProTrrPvalTrrPheAlaAsnLeuSerMetLysAlaTrrPLeuGlyLys 201
   |||||
Db 6543 CTGGGGCTCCCTGGGTGTGTTGCTTCCAACTGATGAAGGCGGTGCTGAGAGTGAAT 6602
OY 201 eArgAspProIleLeuLeuGlnSerLeuLeuSn-----
   |||||
```

Db	6603	CCGGAGCCCTATCTCTGCTCCAGAGCCTGCTGAACGTGAACGCCCACTGTACAGACAGGCTG	6662
Qy	212	-----	212
Db	6663	CCCGAGAGTGGGAAGGGTTGTGTGCTCCACAGAAACAAGTTTCTTCAAAAGAAAGCCTT	6722
Qy	212	-----	212
Db	6723	GGGCCCCCTGAGGCTTTCCGAGAGCCGAGGTGGGGTTGCAGAAATCTTTCCAAACAGCA	6782
Qy	212	-----	212
Db	6783	TCCACAGCCCGAGGTGTCCTTTCAGAGGCCCTCCCTCTTCCAAAGTGTGAGGT	6842
Qy	212	-----	212
Db	6843	CCGTGTTCCCTTTTGTATAGATGAGAGAGCTGAGACAAAGAGTTAGTAGCTTCCCA	6902
Qy	212	-----	212
Db	6903	TGGCCACACAGCCAGGAATGACCAATAGGTACAGGCCCTGTACTGAGAGAGGTGG	6962
Qy	212	-----	212
Db	6963	GGCGGAGCCAGGTGGGGGCGAGGTGTGTTCAGAACCCCATCCCCCTTCTGCCCCC	7022
Qy	213	--GlumetasnThrLeuArpThrIjngysgIyHIsleuTyAlaTyAspTrpIleSerI	232
Db	7023	AGGAGATGAAACCTTGGCTACTGCTGAGTGTGACACCTGTATGCTTACGACTGATTA	7082
Qy	232	LeProLeuValTyThr-----	237
Db	7083	TCCCACTGGTGTATACACAGGTGAGAGTAAAGCTGTGAGGCTGCCTTTTGGGAAACTG	7142
Qy	237	-----	237
Db	7143	AGGCTAGAAGACCAAGGAACAGCTGGGGTGGGAAGGCTCACCTAGAGGCTTAAGTGGC	7202
Qy	237	-----	237
Db	7203	TCCCTGGAGTTGGGTGCACACTTGAAGTTGGGTCTGGACTTGTGAAGTGCAGATTCT	7262
Qy	237	-----	237
Db	7263	AAGATCCAGGCTCTCTGCTGTGCCAGTCCAGTAGAGGAATGTGATTATCCCATATTTA	7322
Qy	237	-----	237
Db	7323	AAGAGAGTTGGCGGCGACAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAAGCTGA	7382
Qy	237	-----	237
Db	7383	GGCAGTGTGATCACCTAGAGGTCAAGAGTTGAGACCAAGCTTGCCAACTGTGAAACCC	7442
Qy	237	-----	237
Db	7443	CATCTTACTGAAATAATACAGAAATTAGCTGTGTGTGTGTGTCAGCGCTGTAAATCCAGCTAC	7502
Qy	237	-----	237
Db	7503	TTTGGAGGCTGAGGACAGGAATCGCTTGAAACCCGGGAGGTGAGGTTGCAGTGAAGTGA	7562
Qy	237	-----	237
Db	7563	GATCATGCACTGTGACCTCGAGCTGGGCGACACAGCAAGACTGTGTCTCAAAACAACAAA	7622
Qy	237	-----	237
Db	7623	CAAAACAACAACAACAAAGGGGTTAAAGAGCCCTTAAGTACATTAAGTGCAGATTC	7682
Qy	237	-----	237
Db	7683	AGAAACAAGGCTTGTGTCTCTGTCTAGACTCCCAAGCCCTTGAGACATCTGATTTCAGG	7742

OY	238	-----	glnValVal	240
Db	7743	GTTCCCACTAGACCCCTTGGTAACACATCTCCTCCCTTCCTCCCTCCACAGTGGTG		7808
OY	241	ThrValAlaValTySerPhePheLeuThcCysLeuValGlyTrpGlnPheLeuAsnPro		2605
Db	7803	ACTGTGGCGGGTACACTCTTCTCCTACTGTGTACTTGGCGGCAGTTTCTGAACCCA		7866
OY	261	AlaIysAlaIyrProGlyHISgluLeuAspLeuValAlaProvalPheThrPheLeuGln		2805
Db	7863	GCCAAAGGCCAACCCCTGGCCCATGAGCTGGACCTGGTGTGGCCCTCTCACGTTCTGACAG		7922
OY	281	PhePhePheTyrrValIGlyTrpLeu-----		288
Db	7923	TTCCTCTTCTAATGTGGCTGGCTGAAGGTGGCCCTCCAGGGCCCTGCTGGGCTGGAG		7988
OY	288	-----	-----	288
Db	7983	CATGGCCAGAGGGGTGATGGCCAGACGTCCTGAGACGAGATGCAGTGTCCAGAAAG		804
OY	288	-----	-----	288
Db	8043	AAGGTCTCACGGGTAGAAAGCACACCAGCGGTGTGGCGCACACCTGTAAITCCCAGCTACT		810
OY	288	-----	-----	288
Db	8103	CGGAGGCGTAGGCGACGAGAAATCGTTGAACCCGGAGGGGAGGTGTGTGATGTAG		816
OY	288	-----	-----	288
Db	8163	ATGTGTCACATGCACTCCACCTGGGCAAAAAGATGAATCACTATCTCAAAAAACAAC		8222
OY	288	-----	-----	288
Db	8223	AACAAACAAACAAAGCCCTAAGATTTCAGAAGCCCCCTTGCCCTTAAAGAAGCAGAGGGCAG		828
OY	288	-----	-----	288
Db	8283	CACCTCTCCTTATTCAGATGSCCTGTGGCGTGTCTGTCTTCTCACTCAAGTGGCTTGT		8344
OY	288	-----	-----	288
Db	8343	CCAGTATTCCTCCACACACAGCAAATACTCGGAAGATGTCTGAATCACACAGATT		8407
OY	288	-----	-----	288
Db	8403	TCTGTCCACCTCTTATCTTTCCTTCCTTCCTGTGTGCCCCACACACTCTCTCCCTTCTCA		8462
OY	288	-----	-----	288
Db	8463	CCTCTCTTATTTTTTGGTAATGGGGGTGAAGTCTGTCTGTGCCCTTCTCTGTCACTG		8522
OY	288	-----	-----	288
Db	8523	TGACACACACACACACACACACACACACACACACACACACATCTTATTCCTCT		8582
OY	288	-----	-----	288
Db	8583	AAATTCCTCCCTGCCCCCAGCTTATCTTGTGTTCTGACATCAAAACMAATCACACTT		8642
OY	288	-----	-----	288
Db	8643	TATGCTGAAATTCCTCAGGGTGGCCAGTGCCCTGCAAGATGTCCCCGTGACCCCTTAAG		8702
OY	288	-----	-----	288
Db	8703	GCAGACGGGTGTACCTCTTCGGGGCTTTGTTAGGGCAATTTAGAGGTGCTATCCAGA		8762
OY	288	-----	-----	288
Db	8763	ATCTGCCACCTAGACTGCCCTTATGTTACAGCCACGTTCAAGTATATACTCTGTTCAT		8822

QY 288 ----- 288
Db 8823 GAATGAATAAATTATGCACTCCAGTAATACATAGGTGAGATPAAAGCAGTACT 8882
QY 288 ----- 288
Db 8883 CAGCCAGATGATACACTCAGGACAGCTGTGGGTTCAGAGGAAGACTGGCTCAGAGA 8942
QY 288 ----- 288
Db 8943 GTTAGAGGGCTGTGTCCAGAAAGTGTGGGTCCCAAGTGTGGGGCTGAGCCCT 9002
QY 288 ----- 288
Db 9003 AAACCTGCTTTGAAGACAGTGTGACAGGAAGGGCTTCATGGGGTGTGAATAGC 9062
QY 288 ----- 288
Db 9063 AGCAGCTGAGGTTTAAAGGGGAAGCTGTGAGAGATTCCTGAGGTTTACGA 9122
QY 289 ----- 288
Db 9123 GCCTCAGCTGTCCCAAGGTGGCAGACAGCTCATCAACCCCTTTGAGAGATGAT 9182
QY 304 ASPhegluThAsnTrPileValAspArgAsnLeu----- 315
Db 9183 GATTTTGACACCACTGATTTGTGACAGGAATTTTCAGTATGGGAGAGAGAGAAA 9242
QY 315 ----- 315
Db 9243 CCATACCATGACCTTCCCAAGTGAACCAAGAGAGACCACCTTGTGAGGA 9302
QY 315 ----- 315
Db 9303 GGCTCAGAGTAAGATCAACCTTCCCTCCTCCCTGACAGCCATTCATCTCA 9362
QY 315 ----- 315
Db 9363 CAGATTCTCACTCAATCTTTGAGGTCGACGAGCACCACATCTCCCATTTACAG 9422
QY 315 ----- 315
Db 9423 CAGGAAACTGAGGTCCAGAGAGAGAGATTCTCCAGTCATCAGCACATCAAG 9482
QY 316 ----- 316
Db 9483 GTCCGCTGGGATGATCTTCTGTGGACTTCTTCTGCTCCCTGGTGAACAGGTGCC 9542
QY 319 uLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetLysTrpAs 9542
Db 9543 GTTGGCTGTGGATGATGACACAGACCTGCTGGATGAGCCGAGACATGACTGAA 9602
QY 339 nLysProGluProGlnProProTyrThrAlaLaserLaglnPheArgAlaSerPh 9602
Db 9603 TAAACCCGAGCACAAGCCCCCTACACAGCTTCCGCCCAAGTTCGTCAGACCTCCT 9662
QY 359 emeGlySerThPheAsnIle----- 366
Db 9663 TATGGGCTCCACCTTCAACATCAGGTGTGCGCAGAGCCAGGGGCTGGTGGAAAGCCC 9722
QY 366 ----- 366
Db 9723 TCCTAGTCAGGGGTCTGCTAGAGAACTTGAATAGACATAGTATGATAGAGTTGC 9782
QY 366 ----- 366
Db 9783 TTCAGTAAGTGCAGACACTGATGCTTTTAAACATTAATTTTTCCTCC 9842
QY 366 ----- 366
Db 9843 CAATATTTCTGTTGTATCCAAAGTTTTCAGATTAATTAAGTACAGGTTCAAGAGAG 9902
QY 366 ----- 366

Db 9903 TAACTGTCCAGAGCCACATAGTACCAAAATGGTGCATTTGTACTGGAAGACAGCTA 9962
QY 366 ----- 366
Db 9963 TGATCAGTATGATGAGTGAACGTTAGACCTGCTCTGTGCATCCAGAACTATGTTCT 10022
QY 366 ----- 366
Db 10023 TTTCTTTTGACAGATATCTGCTGTGTGCGCCAGTTGAGCGCAGTGGGCTATCTT 10082
QY 366 ----- 366
Db 10083 GGCTCAGTCAACCTCCGCTCTGTGGTTCAAGTATTTCTCTCCCTCCGCCAGT 10142
QY 366 ----- 366
Db 10143 AGCTGGGATTACAGGTGCCCAACAACAACGCTGCTAATTTTGTACTTTAGTAGAT 10202
QY 366 ----- 366
Db 10203 GAGGTTACCATGTTGGCCAGGCTGTCTCAACTCTGACAGTAATCTGCCGCTTT 10262
QY 366 ----- 366
Db 10263 GGCTTCCAAATGCTGGAATTAAGTGTCAAAACTATGTTTCTGATTAAGTACATG 10322
QY 366 ----- 366
Db 10323 CTTGATGGGAAGTGAAGTGGGGTCCCTGGATGGGAGAGGAGCAAAAGTCCAGC 10382
QY 366 ----- 366
Db 10383 AGCAGCAGGCCATCAACAGTACCTCTGATTTGTCTTACCAGTAAGAGGCT 10442
QY 366 ----- 366
Db 10443 CAGGCCACCCAGACAGCCAGACTTATCCACATGTGTCCACTTCCGTATTCATCTG 10502
QY 366 ----- 366
Db 10503 AATCCCTTTGAGCTGACAGTGGGCTGAAGGCTATCCAGCTGTCTTCTCCOCAGGA 10562
QY 366 ----- 366
Db 10563 CAACAGAGTTGAAGTGCCTTGGAGAGTGTGGCAGATGTCAAGGTTCAATCAAGG 10622
QY 366 ----- 366
Db 10623 TTTCTTCCACGTAATCCAGTGTCTTCTGCTTGTCTTTTCTTTTAAACG 10682
QY 366 ----- 366
Db 10683 GAGTTTACTTCTTGTGCCCCAGAGCTGAGTGAAGTGAAGGCAATATCTCGCTACTGCAAC 10742
QY 366 ----- 366
Db 10743 CTCGCCCTCCAGATTCAGCAATTCCTGCTCAGCTCCGAGTACTGGATATATA 10802
QY 366 ----- 366
Db 10803 GGTGCGACGACCAAGCCCGCTAATTTTGTATTTTATAGAGACAGTTTCACATGT 10862
QY 366 ----- 366
Db 10863 TGGCAGGCTGTCTCGAACTCTGACCTCAGTGTATCCACCTCCAGCTCCAAAG 10922
QY 366 ----- 366
Db 10923 TGTGGGATTACATGTGTGACCACTGTGCTGCTGTCTTTTAAAGAAACAAATA 10982
QY 366 ----- 366

[illegible]

QY	288	-----	288
Db	8823	GAATGAATAAAATTATGCAACTCCAGGTAAAGATACATGAGGTGAGATAAAGGCAGTGACT	8882
QY	288	-----	288
Db	8883	CAGCCAGTGATACACTCAGGACAGCTGTGGGTGTTTCAGGGAAGACTGGCTCAGAAGA	8942
QY	288	-----	288
Db	8943	GTTAGAGGGCTGTGTCCAGAAGTGTGTGGGTGCCCAAGTGTGGGGGCTGGAGCCCT	9002
QY	288	-----	288
Db	9003	AAACTCTGCTTTGAAGACAGTGTTCAGCAGGAAGGGCTTCATCGGGGTGTGGAATAAGC	9062
QY	288	-----	288
Db	9063	AGCAGCTGAGGTTTAAAGGGGAAGCTGGCTTTGAGGAGTTCCTGCTGAGGGTTTACAGA	9122
QY	289	-----	289
Db	9123	GCCTCACCTGTCCCAAGTGGCAGAGCAGCTCATCAACCCCTTTGGAGAGGATGAT	9182
QY	304	AspPheGluThrAsnTyrIleValAspArgAsnLeu	315
Db	9183	GATTTTGAGACCAACTGGATTGTCACAGGAATTTGCAGGTATGGGGAGAGGAGAGAAA	9242
QY	315	-----	315
Db	9243	CCATACATGGACCTTCCCCAAAGTGAGCCCAAGAGAGAGACCCACTGTTCTGTAGGA	9302
QY	315	-----	315
Db	9303	GGCCTCACAGTGAATGATCAACCTTCCCCCTCCTCCCTGCTGAGCCAGTCACTCA	9362
QY	315	-----	315
Db	9363	CAGGANTCTACCTCAATCTTTGAGGCTGCAGCAGGACCCACTCTCCCATTTACAGG	9422
QY	315	-----	315
Db	9423	CAGGAAACTGAGTCCAGAGAGAGGAGAGATTCTCCAAGTCAATCAGGCACATACAAG	9482
QY	316	-----	316
Db	9483	GTCTGCCCTGGGATGATCTTTCTGTGGGACTTCTTCTGTCCTGTGACCAGGTGTCCCT	9542
QY	319	uLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAs	339
Db	9543	GTGGCTGGGATGAGATGCACAGGACCTGCCTCGATGGAGCCGGACATGTACTGGAA	9602
QY	339	nLysProGluProGlnProTyrThrAlaAlaSerAlaGlnPheArgAlaSerPh	359
Db	9603	TAAGCCGAGCCACAGCCCTACACAGCTGCTTCCGCCAGTTCCTCGAGCTCCTT	9662
QY	359	eMetGlySerThrPheAsnIle	366
Db	9663	TATGGCTCCACCTTCAACATCAGGTGTGGCAGAGCCAGGGGCTGGTGGGAGCCCC	9722
QY	366	-----	366
Db	9723	TCCTAGTGAGGGTCTGCCCTAGGAACCTTAGATAGCACCTAGTAAATGCATACAGGTTGC	9782
QY	366	-----	366
Db	9783	TTCAGTAAGTCTCAGGCACGTACTATGCTCTTTATAAACATTAACATAATTTTCTCCTCC	9842
QY	366	-----	366
Db	9843	CAATAATCTGGTTGTATCCCAAGTTTTCAGATAATTAAGTACAGGTTTCAGAGAGAG	9902
QY	366	-----	366

Db	9903	TAAGTTGTCCAAAGGCCACATAGCTACCAAAATGGTGCAATTTGCTACTCGAAGCAGACCTA	9962
QY	366	-----	366
Db	9963	TGATCAGTGATCCAGTGGAAACGTTAGGACCTGGGCTCTTTGTCTATCCAGAACTATGTTTCT	10022
QY	366	-----	366
Db	10023	TTTCTTTTGTAGACAGTATCTCGCTCTGTGCCCAGGTTGGAGCGCAGTGGCGTGATCTT	10082
QY	366	-----	366
Db	10083	GGCTCACTGCAACCTCCGCCCTCTGGGTTCAAGTGATTCTCTCTTCTCAGCCTCCCCAGT	10142
QY	366	-----	366
Db	10143	AGCTGGATTACAGGTGCCCAACACCACTGCTAATTTTGTACTTTTAGTAGAGAT	10202
QY	366	-----	366
Db	10203	GAGGTTTCACCATGTTGGCCAGGCTGTCTCCAACCTCTGACCAGTAATCTGCCCGCTTT	10262
QY	366	-----	366
Db	10263	GGCCTCCCAAAATGCTGGAAATTATAGGTGTCAAAACTATGTTTCTGATAAGCTACGATG	10322
QY	366	-----	366
Db	10323	CTTGGATGGAAAGTGAAGTGGGGTTCCTCTGGGATGGGGAGGGGAGCAAAAGTCCCAGC	10382
QY	366	-----	366
Db	10383	ASGCAGCCAGGCCATCACAGGTACCTCTCTGAATTGACTTTGTCTACCGAGTAAGGGCT	10442
QY	366	-----	366
Db	10443	CAGGCCACCCACAGCAGCAGCACTTATCCCAACATGGTGTCCCACCTTCTCCCCAGGA	10502
QY	366	-----	366
Db	10503	ATTCCTCTTGAGCTGCAGTGGGTGAAGGGCTATCCCAAGTGTGTCCTGATTCATCTG	10562
QY	366	-----	366
Db	10563	CAACAGAGTTGAAAGTGCTTGGAGAGTGTGGGCACATGTACAGGTTTCATACACAGGG	10622
QY	366	-----	366
Db	10623	TTTCTTCCAGGTATCCAGTGTCTTCTCGCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTT	10682
QY	366	-----	366
Db	10683	GAGTTTCACCTCTTGTGCCCAGAGCTGGAGTGGCATAAATCTCGGCTCACTGCAAC	10742
QY	366	-----	366
Db	10743	CTCCGCTCCAGATTCAAGCAATTCTCTGCTCAGCCTCTGAGTAGCTGGGATTATA	10802
QY	366	-----	366
Db	10803	GGTGCCAGCCACCAAGCCGGCTAATTTTGTATTTTGTAGTAGACAGTTCACCATGT	10862
QY	366	-----	366
Db	10863	TGCCCAGGCTGGTCTCGAACTCTTGACCTCAGGTGATCCACCCTCTCTCAGCCTCCCAAG	10922
QY	366	-----	366
Db	10923	TGCTGGGATTACATGTGTGAGCCACTGTGCTGGCTGTGTCTTTTAAAGAACCAATA	10982
QY	366	-----	366

FT FT /number= 4
FT FT 7228..7461
FT FT /*tag= i
FT FT /number= 4
FT FT 7462..8353
FT FT /*tag= j
FT FT /number= 5
FT FT 8354..8508
FT FT /*tag= k
FT FT /number= 5
FT FT 8509..8896
FT FT /*tag= l
FT FT /number= 6
FT FT 8897..8974
FT FT /*tag= m
FT FT /number= 6
FT FT 8975..9676
FT FT /*tag= n
FT FT /number= 7
FT FT 9677..10032
FT FT /*tag= o
FT FT /number= 7
FT FT 10033..11212
FT FT /*tag= p
FT FT /number= 8
FT FT 11213..11293
FT FT /*tag= q
FT FT /number= 8
FT FT 11294..11556
FT FT /*tag= r
FT FT /number= 9
FT FT 11557..11708
FT FT /*tag= s
FT FT /number= 9
FT FT 11709..13918
FT FT /*tag= t
FT FT /number= 10
FT FT 13919..14557
FT FT /*tag= u
FT FT /number= 10
FT FT 14558..15768
FT FT /*tag= v
FT FT /number= 11
FT FT 15769..16125
FT FT /*tag= w
FT FT /number= 11

PN W09943695-A1.

XX 02-SEP-1999.

XX 22-FEB-1999; 99WO-US03790.

XX 25-FEB-1998; 98US-0075941.

XX 18-DEC-1998; 98US-0112926.

XX (MERI) MERCK & CO INC.

XX (UYUP-) UNIV UPPSALA.

XX Petrukhin K, Caskey CT, Metzker M, Wadelius C;

XX WPI; 1999-540560/45.

XX P-PSDB; AAY29953.

XX Human and mouse polynucleotides encoding CGICE polypeptides

XX Claim 2; Fig 1; 67pp; English.

XX The present sequence represents the human CGICE gene, which when

XX mutated is responsible for Best's macular dystrophy (BMD).

XX Polynucleotides encoding CGICE are useful for diagnosing whether a

XX patient carries a mutation in the CGICE gene. Normal and mutated

XX CGICE proteins are useful for identifying activators and/or inhibitors

CC of these proteins, in order to treat BMD. The CGICE gene offers a
CC simpler and cheaper method of diagnosing BMD without the need for the
CC presence of the patient. The gene may also be useful to discovering
CC the genetic cause of age-related macular dystrophy.

XX Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 21 other;

Alignment Scores:

Pred. No.: 3.09e-150 Length: 16125
Score: 1615.00 Matches: 527
Percent Similarity: 19.86% Conservative: 0
Best Local Similarity: 19.86% Mismatches: 7
Query Match: 51.76% Indels: 2122
DB: 20 Gaps: 7

US-09-622-964-3 (1-585) x AA221226 (1-16125)

Qy 51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluLysLeuThrLeuTyrCysAsp 70
Db 6615 AGGCTGGCCCTCAGGGAAGAACACAGCTAGTGTGAGAACTGACTCTGTATTGGAC 6674
Qy 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
Db 6675 AGCTACATCCAGCTCATCCCCATTCTCTCGTGGTGAGTTCGCCCTTCTGGTGT 6734
Qy 82 ----- 82
Db 6735 CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGCCAGGGAGATCAGGAGAGTGCG 6794
Qy 82 ----- 82
Db 6795 GCAAGGGCTGGGAGGGGGGGAACGCCAGCGAGCTCGGCGCTCTCTGTAGGG 6854
Qy 82 ----- 82
Db 6855 AAAGGTGCGGACTGCAGCCAGAGAACTGAAGTTAGACCTTAGTAAAGACGCTCTGCCGT 6914
Qy 82 ----- 82
Db 6915 TAGCAATGAAGAACCCATTCTGTAGGAAGCGCTGACATCATGTCCCTGGAGCCCTG 6974
Qy 82 ----- 82
Db 6975 CCGGGAGGGAGGGGTCTGGCGGATTTCTGGGACCATTTCTGGGACGAGGGGACCCCGGTGACA 7034
Qy 82 ----- 82
Db 7035 GAACCCCTTGGGGCTCTCGCGCTCCATCGAGGCTCTCCCTGCTCTCGCTCCGAGCGC 7094
Qy 82 ----- 82
Db 7095 CTTCCAGAGGGCTGGGGGTAGCCCGCTCGCAGCAGAAAGCTGGAGAGCCGAGCAT 7154
Qy 82 ----- 82
Db 7155 CGCCGGCGCTGGGCCCTGGGCTCTGGCCAGCGCTGGCCCTCGCCCTCGCCCGCCG 7214
Qy 83 -----GlyPheTyrValThrLeuValValThrArgTyrTrpAsnGlnTyrGlu 98
Db 7215 CCCTCTGCCAGGCTTCTACGTAGCTGGTGTGACCGCTGGTGAACACAGTACGAG 7274
Qy 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
Db 7275 AACCTGCCGTGGCCCGACCGCTCATGAGCTGGTGTGGGTCTCGAAGCAAGGAC 7334
Qy 119 GluGlnSerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
Db 7335 GAGCAAGCGCGCTGCTGGCGCGCAGCTCATCCGCTACCGCAACCTGGGCAACGTC 7394
Qy 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal 158
Db 7395 ATCCGTGGCGAGCGTCAGCACCGCAGTCTACAGCGCTTCCCCCGAGCGCCACCTGGTG 7454


```

Db      13993 GCCTCCTAGGCTGCATGCTCCATGATCACCATCTCCAGGCAAACTCAAGGACCAAAAC 14052
QY      412 euleutrpProLysArgGluSerLeuLeuHISgluLysLeuProLysAsnHISLysAla 432
Db      14053 TACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCCCAAAAACCAAGGACG 14112
QY      432 lALysGlnAsnValArgGlyGlnGluAspAsnLysLAlaTrpLysLeuLysAlaValAsp 452
Db      14113 CCAAAACAGAACGTTAGGGGCCAGGAAGACAAAGGCTTGGAAGCTTAAGGCTTGAGAGC 14172
QY      452 lAPhLysSerGlyProLeuTyrglnArgProGlyTyrTyrSerAlaProGlnTrpPOL 472
Db      14173 CTTAAAGCTGCCCCCAGCTGATTCAGAGGCGGCTACTACAGTCCCCCAGAGCCGCC 14232
QY      472 euSerProThrProMetPhePheProLeuGluProSerAlaProSerLysLeuHISerV 492
Db      14233 TCAGGCCCACTCCCATGTTCTTCCCTTGAAGAACATCAGCCGCCGCAAGCTTCACACTG 14292
QY      492 aLTHrGlyTlAspThrLysAspLysSerLeuLysThrValSerSerGlyLysLys 512
Db      14293 TCACAGGCAATAGACACCAAAAGCTTAAGAGCTGTAGTTGGGGCCCAAGAAA 14352
QY      512 eRPheGluLeuLeuSerGluSerAspGlyAlaLeuMetGluHISProGluValSerGlnV 532
Db      14353 GTTTGAAATGCTCTCAGAGAGCGATGGGCGCTTGATGGAGCACCAAGATATCTCAG 14412
QY      532 aLArgArgLysThrValGluPheAsnLeuThrAspMetProGluLleProGluAsnHISL 552
Db      14413 TGAGAGGAAACAGTGTGAGTTTAACTGACGATATGCAAGAGATCCCGAAATATCACC 14472
QY      552 euLysGluProLeuGluGlnSerProThrAsnLleHISThrLysAspHISMetA 572
Db      14473 TCAAAGAACCTTGGGAACATCAACCAACCAATACACTACACTCAAGATCAACATGG 14532
QY      572 sPRoTyTrrPALaLeuGluAsnArgAspGluLalaHIS 584
Db      14533 ATCCTTATTGGGCTTGGAACAGAGTGTGCTCTCCAC 14570
RESULT 8
AAV99722 standard; cDNA; 1263 BP.
AC      AAV99722;
XX
DT      26-APR-1999 (first entry)
XX
DE      Human adult retina secreted protein bk112_15 cDNA.
XX
KW      Secreted protein; human; retina; bk112_15; ds.
XX
OS      Homo sapiens.
XX
FT      Key Location/Qualifiers
FT      CDS 100..885
FT      /*tag= a
XX
PN      MO9856909-A2.
XX
PD      17-DEC-1998.
XX
PE      08-JUN-1998; 98WO-US11822.
XX
PR      05-JUN-1998; 98US-0092722.
XX
PR      11-JUN-1997; 97US-0873218.
XX
PA      (GENY ) GENETICS INST INC.
XX
PI      Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavallie ER;
PI      McCoy JM, Racie LA, Spaulding V, Treacy M;
DR      MPI; 1999-080899/07.
XX
P-PSDB; AAM95345.
XX

```

```

PT      New polynucleotides encoding secreted human proteins - derived from
PT      human foetal brain, adult testes, foetal kidney, adult thyroid or
PT      adult retina cDNA libraries
XX
PS      Claim 12; Page 71-72; 113pp; English.
XX
CC      This is the nucleotide sequence of cDNA clone bk112_15, which
CC      includes an open reading frame for a 261-amino acid polypeptide
CC      (see AAM95345). The clone was isolated from a human adult retina
CC      cDNA library using methods which are selective for cDNAs encoding
CC      secreted proteins, or was identified as encoding a secreted or
CC      transmembrane protein on the basis of computer analysis of the
CC      amino acid sequence of the encoding protein. Database searches
CC      indicate some sequence similarity to known sequences. The
CC      invention provides cDNA clones (see AAV99721-33) from human adult
CC      thyroid, adult retina, adult testis, foetal kidney and foetal brain
CC      that encode novel secreted proteins (see AAM9344-53). Each clone is
CC      individually available from deposit clone ATCC 98451 (see also
CC      AAV99734-43). The isolated polynucleotides (PNS) and proteins are
CC      predicted to have activities which would make them suitable for
CC      treating, preventing or ameliorating medical conditions in humans
CC      and animals, although no supporting data is given. Suggested
CC      activities include nutritional, cytokine, cell proliferation or
CC      differentiation, immune stimulating (e.g. as vaccines) or immune
CC      suppressing, haematopoiesis regulating, tissue growth,
CC      activin/inhibin, chemotactic/chemokinetic, haemostatic,
CC      thrombolytic, receptor/ligand, anti-inflammatory, cachectin/tumour
CC      invasion suppressor, and tumour inhibition activities. The PNS are
CC      also stated to be useful for gene therapy.
XX
SQ      Sequence 1263 BP; 397 A; 342 C; 272 G; 252 T; 0 other:
XX
Alignment Scores:
Pred. No.: 1
Score: 1571.00 Length: 1263
Percent Similarity: 99.66% Matches: 293
Best Local Similarity: 99.66% Conservative: 0
Query Match: 50.35% Mismatches: 1
DB: 20 Gaps: 0
XX
US-09-622-964-3 (1-585) x AAV99722 (1-1263)
QY      292 GlnGluLeuLleAsnProPheGlyGluAspAspPheGluThrAsnTrpLleVal 311
Db      1 GAGCAGCTCATCAACCCCTTTGGAGGATATGATGTTTGGACCAACCTGATGTGC 60
QY      312 AspArgAsnLeuGlnValSerLeuLeuAlaValAspGluMetHISGlnAsnLeuProArg 331
Db      61 GACAGGAATTTGGCAGGTGTCCCTGTGCTGTGATGATGATGACACAGGACCTCTCGG 120
QY      332 MetGluProAspMetTyrrTrpAsnLysProGluProGlnProTyrrTrpAlaLaser 351
Db      121 ATGAGCGCGGACATGTAATGTAATMACCCGAGGACAGGAGATGCTCACCTGACATCTTCC 180
QY      352 AlaGlnPheArgArgAlaSerPheMetGlySerThrPheAsnLleSerLeuAsnLysGlu 371
Db      181 GCCCAGTTCGCTGAGCGCTCTTATGCGCTTCAACCTTCAACATCAGCTGAACAAAGAG 240
QY      372 GluMetGluPheGlnProAsnGlnGluAspGluGluAspAlaHISAlaGlyLleIleGly 391
Db      241 GAGATGAGTTCACGCGCAATCAGAGGAGAGGAGAGGATGCTCACCTGACATCTTGGC 300
QY      392 ArgPheLeuGlyLeuGlnSerHISAspHISLISProProArgAlaAsnSerArgThrLys 411
Db      301 CGCTTCTAGGCGCTGAGTCCCATGATACCACTCTTCCAGGCGCAAACTCAAGGACCAAA 360
QY      412 LeuLeuTrpProLysArgGluSerLeuLeuHISGlyLysLeuProLysAsnHISLysAla 431
Db      361 CTACTGTGGCCCAAGAGGAATCCCTTCTCCAGAGGCGCTCCCAAAACCAAGAGCA 420
QY      432 AlaLysGlnAsnValArgGlyGlnGluAspAsnLysAlaTrpLysLeuLysAlaValAsp 451
Db      421 GCCAAACAGAACGTTAGGGGCCAGGAAGACAAAGAGGCTGGAAGCTGTGGAC 480

```



```
Db 915 GCTGAGGAGGAGATCAGTCCAGCCAGAGTTCCAGGCTGACGTGCGCTAAGATCGCA 974
QY 160 ----- 160
Db 975 CCGGTGACATCCAACTCGGTGACAGAGCCAGACCCTTCTCTGGAATAATAATATACC 1034
QY 160 ----- 160
Db 1035 CTGCCACATGCTCAGCCAGACAGACCTAGTAGTGTCTGAGAAATTTTGTGTGT 1094
QY 160 ----- 160
Db 1095 GAAAGAAAGAGATGCAAAAGAGTGTCTATAGTGTGAGAGAGTGGCGGCAT 1154
QY 161 -----GlypHemethr 164
Db 1155 CCCTTCGACAGTCTCCACCCAGCCCTTCTTCACTCCACTCTGACAGCTTTATGACT 1214
QY 165 ProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetPheTrpValPro 184
Db 1215 CCGGCAAGAACCAAGAGTGGAGAACTGAGCCTACACAAACATGTTCTGGGTGCC 1274
QY 185 TrpValTrpPheHisAsnLeuSerMetLysAlaTrpLeuGlyGlyArgLysAspPro 204
Db 1275 TGGGTGTGTTTCCCAACTGTCAATGAAGCGTGGCTGGAGTGAATCCGGACCT 1334
QY 205 IleLeuLeuGlnSerLeuLeuAsn ----- 212
Db 1335 ATCTGCTCCAGAGCCGTGTAACGTGAGCCCACTGACAGACAGGGCTGCCAGAGTg 1394
QY 212 ----- 212
Db 1395 GGAAGGCGTGTGTCCACAGAAAGATTTCCTACAAAGAGAGCCTTGGGCCCTGAG 1454
QY 212 ----- 212
Db 1455 GGTCTTCCAGAGCCCTGAGGTGGGGTTGCCAAATCTTTTCCACAGCAATCCAGCCCG 1514
QY 212 ----- 212
Db 1515 AGGTGTCCCTTCTCAGAGGCCCTCCCTCTTCTCCAGTCTGTGAGTCTGTGTCCT 1574
QY 212 ----- 212
Db 1575 TTGTATGATGAGGAAGCTGAGACACAAAGAGTTAGTAGCTTCCATGGCCACACAG 1634
QY 212 ----- 212
Db 1635 CCAGAAATGACCATATGATACCAAGCCCTGTGTAACCTGAGAAAGAGTGGGGCGAGCCCA 1694
QY 213 -----GluMetAsn 215
Db 1695 GGGTGGGGGAGGTGTGTTCAAGAACCCCATCCCTTCTTGTGCCCCCGAGAGATGAAC 1754
QY 216 ThrLeuArgThrGlnCysGlyHisLeuTyralAlaTyraPheTrpHisSerLeuPheVal 235
Db 1755 ACCTTGGCGTACTCACTGTGAGACACCTGATGCTTACGACTGATTAATCCACTGGTG 1814
QY 236 TyrThr ----- 237
Db 1815 TATACACAGTGTAGGATAGCTGTGAGGCTGCCCTTTTGGGAACGTGAGGCTAGAAG 1874
QY 237 ----- 237
Db 1875 ACCAAGAAAGACAGCTGGGGTGGGAAGGCTCACTAGAGAGCTAACTGCTCCCTGGAG 1934
QY 237 ----- 237
Db 1935 TTGGGTCCACACTTTGAAGTTGGTCTGTGACTTTGAAGTCCCAAGTTCTAAGATCCAGG 1994
QY 237 ----- 237

Db 1995 CTCCTGCTGGCCAGTCCAGTAGAGCAATGTGATTATCCCATATTAAAGAGGTTG 2054
QY 237 ----- 237
Db 2055 GCCGGGTGAGTGGCTCATGCTGTGTAATCCAGCACTTTGGGAAGCTGAGAGGTTGAT 2114
QY 237 ----- 237
Db 2115 CACCTGAGTTCAGAGTTCGAGACAGCCTGGCCAAACATGTGAACCCCATCTACTG 2174
QY 237 ----- 237
Db 2175 AAAATACAGATTAGTGTGTGTGTGTCAGCGCTGTAAATCCAGCTACTTGGAGCTG 2234
QY 237 ----- 237
Db 2235 AGCAGAGAAATCGTTGAACCGGGAGGTGGAGTTGCAGTGAGCTGATCATGCGAC 2294
QY 237 ----- 237
Db 2295 TGCATTCAGCCTGGGCGACAGACAAAGACTGTGTCAACAAACAAACAAACAA 2354
QY 237 ----- 237
Db 2355 ACAAGGGGTTAAGAGAGCCCTAAGTCAATAGTGTGCAAGTCAAGAACAGGCTTGG 2414
QY 237 ----- 237
Db 2415 TCTCTGTCTCAGACTCCAGCCCTGGAGCATTCGATTTCAGGGTTCCACCTAGCCC 2474
QY 238 -----GlnValValThrValAlaValLysSerPhePhe 248
Db 2475 TTGTGACCAATCCTCTCTCTCTCTCTCCAGGTGGTGTGCTGAGTGTGAGTCTTTC 2534
QY 249 LeuThrCysLeuValGlyArgLysPheLeuAsnProAlaLysAlaTyraProGlyHisGlu 268
Db 2535 CTGACTTGTCTGTGTGGGGGAGTTTCTGAACCCAGGCAAGCCTTACCTGGCCATGAG 2594
QY 269 LeuAspLeuValValProValPheThrPheLeuGlnPhePhePheTyraGlyTrpLeu 288
Db 2595 CTGGACCTGTGTGTGCTCCCTTCTCAGCTTCTGAGTCTCTTCTTATGTTGGCTGTG 2654
QY 288 ----- 288
Db 2655 AAGTGGGCTCTCCAGGGCCCTGTGCGTGGAGCATGGCCAGAGGGTTCATGCGCAG 2714
QY 288 ----- 288
Db 2715 CAGTGTCTGAGAGAGAGTGCAGTGTGTCAGAAAGAGTCTCAGGGTAGAAAGCAGC 2774
QY 288 ----- 288
Db 2775 CAGCGGTGTGGCGACACCTGTATATCCACTACTCGGAGGCTGAGGCGAGAAATCG 2834
QY 288 ----- 288
Db 2835 CTTGAACCGGAGAGCGGAGGTGTGTGATGTGATGATGTCGACACTGACACTCGACCTG 2894
QY 288 ----- 288
Db 2895 GGCAAAGAAATGAATCTATCTATCTCAAAAACAAACAACAACAACAAAGCCTAAGGT 2954
QY 288 ----- 288
Db 2955 TCAGAGCCCTGCTTGAAGGAGCATGGGACCACTCTCTTATTCAGATGCTCT 3014
QY 288 ----- 288
Db 3015 GTTGGGTGTCTTGTCTCAGCTCAAGTGTGCTCAGGTATTCCTCCACACAGCAGC 3074
QY 288 ----- 288
Db 3075 CAATACCTCGAAGAGATGTTCTGATATCAGACAGTTTCTCTCCACCTTATCTTCTCT 3134
```

QY	316	-----GlnValSerLeuLeuAlaValAspGlu	324
Db	4214	TCCTTCGTGGAGCTTCTCTGTCCTCCGTGACACAGGTGTCCTGTTGGCTGTGATGAG	427
QY	325	MetHisGlnAspLeuProArgMetGluProAspMetGluTrpAsnLysProGluProGln	344
Db	4274	ATCCACAGAGACTGGCTCGAGGAGGAGCGAGCATGTACTCGAATAAAGCCGAGCCACAG	433
QY	345	ProProtyrThrAlaSerAlaGlnPheArgAlaSerPheMetLysSerThrPhe	364
Db	4334	CCCCCTACACAGCTGT -TCCGCCAGTTCCTCGAGCTCCTTAATGGGCTCACTTC	439
QY	365	AsnIle-----	
Db	4393	AACATCAGGTGTGCCAGAGCCAGGGGCTGGTGGGAAGCCCTCTCAATGACAGGGTTC	445
QY	366	-----	
Db	4453	TGCCTAGAAGCTTGAATAGACATGTAATGATACAGTGTGCTCAGTAAGTGTACAG	4512
QY	366	-----	
Db	4513	CACGTACTATGCTCTTTATTAACATTAACATTTTTCCTCCCAATATTCGTGTTG	4572
QY	366	-----	
Db	4573	TTATCCCAAGTTTCAATATTAATAAGTACAGGTTGAGAGAGTAAGTGTCCAAAGCC	4632
QY	366	-----	
Db	4633	ACATAGCTAACAAATGTCATTTGCTACTCGAAGACAGCCTGTATGATGATGATGACAGT	4692
QY	366	-----	
Db	4693	GGAGCGTAGAGACCTGGCTCTTGTCATCCAGAACATGTTCTTTCTTTTGGAGACAG	4752
QY	366	-----	
Db	4753	TATCTGCTCTGTCGCCCGCAGGTTGGAGCGCAGTGGCGTGAATCTTGCGTACATGCACCTC	4812
QY	366	-----	
Db	4813	CGCCTCTGGGTTCAAGTATCTCTCGCTTCAGCCTCCACAGTAGCTGGGATTAACAGT	4872
QY	366	-----	
Db	4873	GCCCAACAACCAACTGGCTAATTTTGTACTTTAGTAGAGATGAGGTTTCAACATGTT	4932
QY	366	-----	
Db	4933	GGCAGAGCTGGTCTCAACTCTGACCAAGTAATCTGCCGCTTGGCCTCCAAATATGCT	4992
QY	366	-----	
Db	4993	GGAATTAAGTGTCAAAACTATGTTTCTGATAAGCTACGATGCTTGGATGGAAAGTG	5052
QY	366	-----	
Db	5053	AAGTGGGTTTCCTGGGATGGGGAGGGCAGCAAAATCCAGACAGGACAGGCCATC	5112
QY	366	-----	
Db	5113	ACAGTAGCTCTGAATGACTTGTCTTACGAGAGTAAAGGCTCAGGCCACACACAGCA	5172
QY	366	-----	
Db	5173	GCCAGACTTATCCCAATGATGCCACTTCCCTGATTCATCTGAATCCCTCTTGAAGTG	5232
QY	366	-----	
Db	5233	CAGTGGGCTGAAGGGCTATCCAGCTGTCTTCTCCCAAGCAACAGAGTTGAAGT	5292
QY	366	-----	

Dh 5293 GCCCTGGAGAGTGTGGGCACATGTCAGGGTTCACTCAAGGTTTCTTCCAGGTAATC 5352
Qy 366 ----- 366
Db 5353 CAGTGTGTCTTCGCTTGTCTTCTTTCTTTTAAAGGAGTTTCACTCTGT 5412
Qy 366 ----- 366
Db 5413 GCCCAGAGCTGGAGTGCAGTGGCATATCTGGGCTCAGTCAACCTCCGCCCTCCAGATT 5472
Qy 366 ----- 366
Db 5473 CAAGCAATTTCTGCTCCCTCAGCTCTGATAGTGGATTATAGTGCACAGCCCAAG 5532
Qy 366 ----- 366
Db 5533 CCCGGCTAATTTTGTATTTTATAGTACAGTTCACATGTTGGCCAGGCTGGTCTC 5592
Qy 366 ----- 366
Db 5593 GAACCTCTGACCTCAGTGCATCAACCTCTCAGCCCTCCAAAGTGTGGATTACATGT 5652
Qy 366 ----- 366
Db 5653 GTGAGCCACTGTGCTGCTGCTGCTTTTAAAGAACCAATATCTCTAGACTGCAA 5712
Qy 366 ----- 366
Db 5713 TCGAGTTTAACTACAGTCTATAGTATGAGGAGGTTGGAGAGTGTATCAATGAA 5772
Qy 366 ----- 366
Db 5773 GCGTGGAGGCTTCTTAGGTGAGAAACATTTCTGAGAGTACTTTGAGCCCTACATGT 5832
Qy 366 ----- 366
Db 5833 CTGTACCCACAGACTGAAGTTTGTAGGGGTGGGAGGCTGTAACAGACGATGAA 5892
Qy 366 ----- 366
Db 5893 GCATAGACCTTGTCTCCAGAGAAATGACAAATTTATGAGAGGAGCTCAACCCAGCCTCA 5952
Qy 366 ----- 366
Db 5953 AACTCTGGATACAGGATCAAAAGTACTGATGTCAGAAAGGAGACAGAAATGGAACAC 6012
Qy 366 ----- 366
Db 6013 AGTCATCTTGTCTGCTGGAGGCGGCTTCACAGCTGGTCTGAGCTGAGCCATGGAAC 6072
Qy 366 ----- 366
Db 6073 ATGGGAAGATCTGAACCTTGGGCAAGGGCAGGCCATCTCTGTAGATAGCTTTCC 6132
Qy 366 ----- 366
Db 6133 TTGCAGGGTAAGGTCTGGGGCTCCGGGATGCTGTGTAGAGTAATTTCTCTT 6192
Qy 366 ----- 366
Db 6193 TGTGATGTACATCCAGTTGGAACCAAAATTCCTGGCATTTGCCAGAGTCACTCATGGG 6252
Qy 366 ----- 366
Db 6253 CCTCATCTGAACACATCATGCCAGGGCAGCAGTGTCTTCTGACTGCTGTGAGGGGGTT 6312
Qy 366 ----- 366
Db 6313 TTACAGGGAAAGTGAATGATGAGAGGCTTTACAGCCAGGGGGGGGTTCGGGGGG 6372
Qy 366 ----- 366

Dh 6373 TTGATGTTAACTCTGGTCAAGAGGAATCAACAAGATGAGTGAAGTGGCCGTGAG 6432
Qy 366 ----- 366
Db 6433 GGATCACCAGGAGGTACAGAGACATCAGAGAGAGGTGAGAGCTGGGGCATGTGAGGA 6492
Qy 366 ----- 366
Db 6493 AGACGGTGTGGCTTGGCTTGGCCAACTGAGAGAGAGCGGGGTAAGGAGAGTA 6552
Qy 367 ----- Ser 367
Db 6553 AGCCAGGTGTGGTCTTGTCCAGCTGCTACCCCTGCATCTCTCTTTCTTCAGC 6612
Qy 368 LeuAsnLysGluGluMetGluPheGluProAsnGluLysAspGluAspAlaHisAla 387
Db 6613 CTGAACAAGAGGAGATGGAGTTCAGCCCAATCAGAGAGCAGAGAGATGT-CACGCT 6671
Qy 388 GlyLeuLeuGlyArgPheLeuGlyLeuGlnSerHisAspHisIleProPheArgAlaAsn 407
Db 6672 GGCATCTTGGCCGCTTCTTAGGCTGCAAGTCCCATGATCACCATCTCTCCAGGCAAC 6731
Qy 408 SerArgThrLysLeuLeuTyrProLysArgGluSerLeuLeuHisGluLysProLys 427
Db 6732 TCAAGGACCAACTACTGTGGCCCAAGAGGAAATCCCTCTCCACGA-GGCCCTGCCAA 6790
Qy 428 AsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGluAspAsnLysAlaTyrLysLeu 447
Db 6791 AACCCAGAGGACCCCAACAGAACGTTAGGGGCAAGAAACAAAGAGCTTGGAAGCTT 6850
Qy 448 LysAlaValAspAlaPheLysSerGlyProLeuTyrGlnArgProGlyTyrTyrSerAla 467
Db 6851 AAGGCTGTGGAGCCCTTCAAGTGTGCCCACTATACAGAGCCAGCTACTACAGTGC 6910
Qy 468 ProGlnThrProLeuSerProThrPrometPhePheProLeuGluProSerAlaProSer 487
Db 6911 CCACAGACGCGCTCAGCCCACTCCCAAGTCTTCCCTTGAACATCACAGCGCGCTCA 6970
Qy 488 LysLeuHisSerValThrGlyIleAspThrLysAspLysSerLeuLysArgValSerSer 507
Db 6971 AAGCTTCACAGTGTCTCAGAGCATAGACACCAACAAAGCTTAAAGCTGTGAGTCT 7030
Qy 508 GlyAlaLysLysSerPheGluLeuLeuSerGluSerAspGlyAlaLeuMetGluHisPro 527
Db 7031 GGGGCCCAAGAAAAGTTGAAATTCCTCAGAGAGGATGGGGCTTGATGAGACACCA 7090
Qy 528 GluValSerGluValArg 533
Db 7091 GAAGTATCTCAAGTGAAG 7108
RESULT 10
ABL89697
ID ABL89697 standard; cDNA; 1717 BP.
XX ABL89697;
AC
XX
XX 24-MAY-2002 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 259.
DE
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antinflammatory; antilucer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
OS Homo sapiens.
XX
XX
XX WO200190304-A2.
PN
XX 29-NOV-2001.
PD
XX
XX 18-MAY-2001; 2001MO-US16450.
PF

XX 19-MAY-2000; 2000US-205515P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Blirze CE, Rosen CA;
 XX WPI: 2002-122018/16.
 DR P-PSDB: ABB89288.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders -
 XX
 PS Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 9 other;
 Alignment Scores:
 Pred. No.: 1,54e-94 Length: 1717
 Score: 1045.00 Matches: 200
 Percent Similarity: 97.56% Conservative: 0
 Best Local Similarity: 97.56% Mismatches: 5
 Query Match: 33,49% Indels: 1
 DB: 24 Gaps: 0
 US-09-622-964-3 (1-585) x ABL89697 (1-1717)
 QY 380 GUAAGPGLUGLUAAPLAHISALAGLLEGLYATGPHLEUGLYLEUGLINSERHS 399
 DB 1 GAGGACGAGGAGGATGCTACGCTGATCATTTGGCCCTCTAGGCTGACATCCAT 60
 QY 400 AAPHISHAPROFROATGALAAASNSERATGTHRLYSLEULEUTRPPOLYARGLUSER 419
 DB 61 GATCCACCTTCCTCCAGGCAAACTCAAGACAACTACTGTGGCCCAAGGGAATCC 120
 QY 420 LEULEUHSGLUGLYLEUPROLYASNSHISLALALYSGLINVAIAIGLYLN 439
 DB 121 CTCTCCACGAGGCTGACCAAWAACCCACAAAGCAGCAAGAAAGAGGAGGAG 179
 QY 440 GLUASPAALYSALATPLYSLEULYSALVALASPLAPHELYSSEGLYPROLEUTYR 459
 DB 180 GAAGACAAACAAGGCTGGAAGCTTAAGGCTGTGACCCCTTCAAGTCTGCCACTGTAT 239
 QY 460 GLATGPRGGLYTYTYRYSERIALAPROGLINTHPROLEUSERPROTHPROMETPHEPHE 479
 DB 240 CAAGGCGCAGGCTACTACAGTCCACAGACACGCCCTCCACTCCCATGTTCTTC 299
 QY 480 PROLEUGLUPROSERIALAPROSERLYSLEUHSSEVALTHRLYILEASPHRLYSASP 499
 DB 300 CCCCTAGAACATACACGCCGCTCAAGCTTCACAGTGTACAGGATAGACACCAAGAC 359
 QY 500 LYSERLEULYSRTHVALSERSERGLYALALYSLSERPHGLULEULEUSERGLUSER 519

DB 360 AAAAGCTTAAAGACTGAGAGTCTGGGCGCCAGAAAAGTTTGAATTGCTCTCAGAGAC 419
 QY 520 ASPLGLYALALEUWETGLUHSPPROGLUVALSERGLINVALARGARGLYSRTHVALGLUPHE 539
 DB 420 GATGGGGCTTGTGATGAGCACCCAGAGTATCTCACTGAGGAGAAACCGTGGAGTTT 479
 QY 540 ASNLEUTHRASPECTPGIULIEPROGLUASNSHISLULYSGLUPROLEUGLINSER 559
 DB 480 AACCTACGATATGCTGACAGATCCCGAAAATCACCTCAAGAACCTTTGGAACATCA 539
 QY 560 PROTHASNLIEHSRTHRLYSASPHISMEASPPROTHRTPALALEUGLUSAN 579
 DB 540 CCAMCCACATACACACTACACTCAAGATCATGATCTTATTTGGGCTTGGAAAC 599
 QY 580 ARGASPLALAHIS 584
 DB 600 AGGTCTGCTCTCCAC 614
 RESULT 11
 ABL10793
 ID ABL10793 standard; cDNA; 2861 BP.
 XX
 AC ABL10793;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001MO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 XX
 DR P-PSDB; ABB66690.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 26861; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences and pharmaceutical drugs. The invention
 CC sequences (ABL10840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 2861 BP; 761 A; 701 C; 723 G; 676 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,34e-90 Length: 2861
 Score: 1009.00 Matches: 250

Percent Similarity:	48.90%	Conservative:	83
Best Local Similarity:	36.71%	Mismatches:	216
Query Match:	32.34%	Indels:	132
DB:	23	Gaps:	18

Qy	1	MethrIlethrYrThrSerGlnValAlaAsnAlaArg--LeuGlySerPheSerArg	19
Db	415	ATGCAATTAACGTACACAGGTAAGGGCCATTGCGGGCTTGGCTGTTTCTCAAA	474
Qy	20	LeuLeuLeuCystrPAAGGlySerIleTyrIlyIleuLeuYrGlyGlnPheIlePhe	39
Db	475	TTGGCTGCAGATGCGCAGGACGACATTTCACAACTGGTTGGCTAGATCTTGGCCCTC	534
Qy	40	LeuLeuCystrYrIleIleIleArgPheIleTyrArgLeuAlaLeuThrGlnGlnGln	59
Db	535	TTGACCACTTACTATGATGGATCAACATGGTATGCTTGGCCCTCAACCCCGACAAA	594
Qy	60	LeuMetPheGlnLysLeuThrLeuIlyCysAspSerYrIleGlnLeuIleProIleSer	79
Db	595	GAACCTTTGAGGCCATTTCTTCACTGCTGGAATGATACAGAAACATCAACCCCTGCC	654
Qy	80	PheValLeuGlyPheTyrValThrLeuValValThrArgTrpAsnGlnTyrGlnAsn	99
Db	655	TTTCGTCCTTGTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	714
Qy	100	LeuProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGln	119
Db	715	ATTCCCTGGCGAATCCCATCCGCGGTTTGTACGTCGATGATGATGATGATGATGATG	774
Qy	120	GlnSerArgLeuLeuArgTrpThrLeuIleIleArgTyrAlaAsnLeuGlnValLeuIle	139
Db	775	CGAGAGCGCATGATGAGCGAACAATATGATGATGATGATGATGATGATGATGATG	834
Qy	140	LeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGln	159
Db	835	CTGGCGAATGTTTCCGCCAGGGGAAGACGCTTCCCGCGCTAATATATCTGGTGGA	894
Qy	160	AlaGlyPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeu-----SerLeuPro	177
Db	895	GGGGCTCTCTAATGACAAATGAAGAAGACATCATGACATGACATGACAAAGGCTTCCC	954
Qy	178	His----AsnMetPheTrpValProTrpValThrPheAlaAsnLeuSerMetLysAlaTrp	196
Db	955	AGACCTTCGAACCACTGGCTGCCATCCGTTTGGGCGCAGATATATATACAGGGCCANA	1014
Qy	197	LeuGlyGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThr	216
Db	1015	AAGGAAGGCGCATTCGTCATATTTTGGTGTGGAAGACCATATGATGATGATGATGATG	1074
Qy	217	LeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyr	236
Db	1075	TTTCGTGCTCACTGTGGAGCTCTCTCAACGCTCGATACCATTAAGTGTACCTTGCTGAC	1134
Qy	237	ThrGlnValValThrValAlaValArgSerPhePheLeuThrCysLeuValGlyArgGln	256
Db	1135	AACCAAGGCGACCTCGCGGTGATTCGTAATCTTACTGCTGATGGCTGACAA	1194
Qy	257	-----PheLeuAsnProAlaLysAlaLysPro	265
Db	1195	TGAGCCGATGGCAAGTGTGGGCATACCACTACTACGTACAC-----	1254
Qy	266	GlyHisGlnLeuAspLeuValValProValPheThrPheLeuGlnPhePhePheTyrVal	285
Db	1237	-----AAGTGGATCTACTTCTTCCTTATTTACACGCGCGAGCTTCTTCTACATG	1290
Qy	286	GlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGlnAspAspAspPhe	305
Db	1291	GGTTGGCTCAAGGTGGCGAAGTCCGCTGTAATCATTTGGCGAAGACGATGATGATTTT	1350
Qy	306	GlnThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeuAlaValAspGlnMet	325

DB	1351	GAGGTCACTGGATGGTGGATGCGCAATCTTCAGAGTGTCTATGTGATGCTGAGAGATG	1410
QY	326	HISGLASPRLEUPROALGNETGLUPROASMETTYTTPRASNLYSPROGLUPRO----	GLN 344
DB	1411	CACATGACCATCCCGAGAGCTGTAAAGAGTCAGTACTGGAGCAGAGTCTCCCAACAG	1470
QY	345	PROPTGYTHRALALASERLAGINPREADGALGALASERPHEMETGLYSETHPHE	364
DB	1471	CTGGCCCTACACATTAAGTCTGCCAGACGATCCGGGAAATCATCCAGAGCCCTCCAGTCC	1530
QY	365	ASNILESERLEUASNLYS-----GLUGLIMETGLUPHGLIN	376
DB	1531	AAAGATGAGGTGGCCCAAGAAATGCGGAGATGCGATGCGACATGATGTGCTCGGATGAT	1590
QY	377	PROAGNGLINLUPROGLINLUPRNLINLS-----	386
DB	1591	GAATATGGCCGATGATCCAGTGGCATTCATCTTCTCAGCTGGAATAATGCCAAAATCGGCTG	1650
QY	387	-----ALA 387	
DB	1651	GATTCCTCGCCCTCGGTGAGCTTTCGGGAACCTGATCCGGGATGATACGGTGGC	1710
QY	388	GLYLIELGLLYARGPHELEUGLY-----LEU 396	
DB	1711	TCGGCCCTCAACAGCTTTCCTGAGCCGCGAGCATGACAGCCGGAGTTCAGGCCCTG	1770
QY	397	GLINSETHASPRHSHASPRGALALANSEARGTHLYSLEULEUPROLYS	416
DB	1771	CAGGATCAGCCCTCAATAATCCCGCGCAGCCGACCTGGGAGTTATGCGGCGCTG	1830
QY	417	ARGGLUSERLEUENHISGLUGLYLEUPROLYASNHSYSLALALALYSGLINLVAL	436
DB	1831	GRAGATGCGGCTACATCGCCGCGGAAACACCTGGCATGCTTGAATTAACGAGCAAGTG	1890
QY	437	ARGGLYGLINLUPRASNLYSALA-----TRPLYSLEULYSALVALAASP	451
DB	1891	ATCGAGAGGTGGAGCAACAGCGCATACATCATCTGAGAGCAATGATCCAGCTGCC	1950
QY	452	-----ALAPHEUSSERGLYPROLEUTYR	459
DB	1951	AATGTCAATGACATTTCTGGCAACAACTCGTCGGAGACTGGAACCTCTGGAACCGCTGAG	2010
QY	460	GLINATGPROGLYTYTYSERIALAPROGLINTHPROLEUSERPRO-----	474
DB	2011	CCACCACCGCGCCACTCGGAACCGGTGAGATCCCGTCAGCTCGGCGCTCATCAATCGG	2070
QY	475	-----THYPROMEPHPPHRO-----LEUGLIN 482	
DB	2071	GCCCAATCCCACTACAGAACCCAACTATTTCACCTCGCGAGATGATGCACGTGCTCACT	2130
QY	483	PROSERIALAPROSER-----LYSLEUHSERVALTHRGY-----	494
DB	2131	ACTTCGATCTCGTGGCGGAGATCCCTGCTCTGTGTAATCAGCAGCACTGCACCCACT	2190
QY	495	-----LIEASPTNLYASPLYSERLEULYSTHVALSERSELYLALALYSYLS---	511
DB	2191	TCGGCAGTGGCGGAGACTCCAAAGTCTCTATACGATTCACAAAAGGCGCCGCGAGAG	2250
QY	512	-----SERPHEGLIULEULEUSERGLUSERASP-----GLYALALEUMETGLINLS	526
DB	2251	ACAAGTGAAGACAGGACCTGAGATGCTCCACGAGATCTACGCGAGATCGGCACTGGCAG	2310
QY	527	PROGLINLVALSERGINLVALARGTGLYTHVALGINLPHASNLEUTHASPMETPROLI	546
DB	2311	CCCGAAGACGAGGAGC-----GATGACTTCGATAG	2340
QY	547	ILERPGLIASNHSLEULYSGIUPROLEUGLINSERPPTROTHASNILEHISTHTRHR	566
DB	2341	CTGAAGCGCGAAGCGGAGAGAACTGATGCGACACAAAGATCTGCGCCAGAACT	2400
QY	567	LEU 567	
DB	2401	ATT 2403	

QY 403 ProProArgAlaAsnSerArgThrIysLeuLeuTrpProlys 416
 Db 1222 -----GCCCGCCGGAGTTCAGTTGGAGCAAA 1248
 RESULT 13
 ID ABL10792/c standard; cDNA; 10760 BP.
 XX ABL10792;
 AC ABL10792;
 XX 26-MAR-2002 (first entry)
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26858.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI: 2001-655686/75.
 DR P-PSDB; ABB66689.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 26858; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB163051), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 10760 BP; 3232 A; 2079 C; 2173 G; 3276 T; 0 other;
 Alignment Scores:
 Pred. No.: 6.02e-74 Length: 10760
 Score: 851.50 Matches: 209
 Percent Similarity: 44.23% Conservative: 71
 Best Local Similarity: 33.02% Mismatches: 143
 Query Match: 27.29% Indels: 210
 DB: 23 Gaps: 13
 US-09-622-964-3 (1-585) x ABL10792 (1-10760)
 QY 21 LeuLeuCytyrArgIleSerIleTyrlsLeuLeuTyrglyLuhPheLeuIlePheLeu 40
 Db 5539 CTTTATGAGTGGGAGAGACATTACAACTGGTTGGCTATGCTTCTGGCTTCTG 5480
 QY 41 LeuCytyrTyrlleIleArgPheIleTyrgArgLeuAlaLeuThrGluGluGlnGln--- 59
 Db 5479 ACCATTACTATGCGATCAACATGCTATCGCTTGGCTCAACCCGACACAAAAGAG 5420

QY 59 -----
 Db 5419 TGAGTATAAAGCTCTAGAGAGATGAGATTGGTGCCTACATTAAACATA 5360
 QY 59 -----
 Db 5359 TTTATTGAATTGAATATTATTATTTTGAATTATCTATGTTCCATTCAATTCAGAA 5300
 QY 59 -----
 Db 5299 AGTGCATGGGCTCTAAGTCAATACATGATGATTAATTACTTAAATACAGACAGTA 5240
 QY 59 -----
 Db 5239 GAAAGTTTACTCAAAACATCTGAAATAGCTTTGACAACTTCACAAATGGTGTAAAT 5180
 QY 59 -----
 Db 5179 AAAATTAGCTGATGCTACAAATACATACATATATGATGACAGATCTTCTCCACGA 5120
 QY 60 -----LeuMetPheGluLys 64
 Db 5119 TCTTGTCTCTTTTAAAGCCCTTACTAATATGCAACTCTTTCAGAACCTTTGAGCC 5060
 QY 65 LeuThrLeuTyrcysAspSerTyrlleGlnLeuIleProIleSerPheValLeuGlyPhe 84
 Db 5059 ATTGTCAGTACGTGTACTTACTACAGAACTATACCCCTGCTCTGCTGCTGCTTC 5000
 QY 85 TyrValThrLeuValIleThrArgTyrTrpAsnGlnTyrgLysLeuPheProIleProAsp 104
 Db 4999 TATGTATGATTTGTATGATGACCCGTGTGTAACATGACACCTTCATCTCCGCGCAT 4940
 QY 105 ArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGlnSerArgLeu 124
 Db 4939 CCCATGCCCTGTTGTACAGCTCAATGTCCATGACGAGATGAGCAGAGCATGATG 4880
 QY 125 ArgArgThrLeuLeuArgTyrlleAlaAsnLeuGlyAsnValLeuLeuLeuArgSerValSer 144
 Db 4879 AGCGCAACATATATGCGATATGTGTGCTTGTGCTGACGATGGCTCGGCGGAATGTTTCG 4820
 QY 145 ThrAlaValTyrlsArgPheProSerAlaGlnHisLeuValGlnAlaGlyPheMetThr 164
 Db 4819 CCGAGGTGAAGAGCGTTCCCGGCGCTAAATATCTGTGTGAAGGGGTCTGCTAAAT 4760
 QY 165 ProAlaGlnHisGlnLeuGluLysLeu-----SerLeuProHis---AsnMetPhe 181
 Db 4759 GACAAATGAAGAGCATCATCGAGACCATGAACAGGCCCTTCCAGACCTTGAAACAC 4700
 QY 182 TrpValProTrpValTrpPheAlaAsnLeuSerMetIysAlaTrpLeuGlyGlyArgIle 201
 Db 4699 TGGCTGCCCATCTTGTGGCTGCCATATATTAACAGGCCGAGAAAGAGTGCAT 4640
 QY 202 ArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGlnCys 221
 Db 4639 CGTGATGATTTGCTGTGAAGACCATCATCATGATGACTAAATAGTTTCTGTGCTAGTGT 4580
 QY 222 GlyHisLeuTyrlleAlaTyrlsArgTrpIleSerIleProLeuValTyrlsGlnValAlaThr 241
 Db 4579 GGAAGCTCATGAGCTGATACCATGATGATGATGATGATGATGATGATGATGATGATGAT 4520
 QY 242 ValAlaValTyrlsSerPhePheLeuThrCysLeuValGlyArgGln----- 256
 Db 4519 CTGGGCGGTATTCGTACTTCTTACTCTGTCGATGGGTCAACAAATGAGCCATGGCAAG 4460
 QY 257 -----PheLeuAsnProAlaLysAlaTyrlsProGlyHisGluLeuAsp 270
 Db 4459 GTGGTGGCAATACACATACCTGAAAC-----AAGGTGGAT 4424
 QY 271 LeuValValProValPheThrPheLeuGlnPhePhePheTyrlleGlyTrpLeuLysVal 290
 Db 4423 CTATACTCTCTGTATTTCAACAGCTGCAAGTCTTCTTCTACATGGGTGGCTCAAGGTG 4364
 QY 291 AlaGluGlnLeuIleAsnProPheGlyLysAspAspAspPheGlu----- 306

```

Db 4363 GCGAGTCGCTGATAATTCATTGGCGAGACGATGATTTGGAGGTACGATTAGA 4304
QY 307 -----Thrasylleval 311
Db 4303 ACACATTTGTTAGTTCTCCGAAAAATCAATAAATATATCTCCAGGTCACATGGATGTG 4244
QY 312 AsparaginLeuGlnValSerLeuLeuValAspGluMetHisGlnAspLeuProArg 331
Db 4243 GATGCAATCTTCAGGTGCTCTATCTGATGCTGACGAGATGCACATGACCATCCGAG 4184
QY 332 MetGluProAspMetTyrTrpAsnLysProGluPro--GlnProTyrThrAlaAla 350
Db 4183 CTGTAAAGATTCAGTACTGGAGACGAGTGTCCCAACGAGTCCCTACACATAGCT 4124
QY 351 SerAlaGlnPheArgAlaSerPheMetGlySerThrPheAsnLysSerLeuAsnLys 370
Db 4123 GCGCAACGATTCGGGGAAATCATCCAGACCCCTCCCAAGATCGAGGTGCCAAG 4064
QY 371 -----GluGluMetGluPheGlnPro 377
Db 4063 AATGCGGCATGCCATTCACATATGTCGTCGATCGATGATAATGGTTGTAATGCT 4004
QY 378 AsnGlnGluAspGluGlnAspAlaHisAla----- 387
Db 4003 AATGCGATCCGCGCCCGCTCGGATGCGAAAGCCCGAACCCGGTTACGGGTCAAGTT 3944
QY 388 GlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAspHisHis-ProProArgAlaAs 407
Db 3943 CAAGTTCGCGGTGCGGACACGCGCGTCCCAACCGTACTACTGATTTTCCCAAGGCTCG 3884
QY 407 nSerArgThrLysLeu-----LeuTrpProLysArgGluSerLeuLeuHisGlu 424
Db 3883 CCAAGTAGAGAGGTTGAATCAAAACGTTGGCCG-----CATCCCTAGCCTTACGAAG 3830
QY 424 yLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGluAspAsnLysAl 444
Db 3829 CATTCGCACACAGACATCATCGGTTAACAAAT----- 3796
QY 444 aTrpLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrGlnArgProGlyTyr 464
Db 3795 -----TCATCATTCCTCCCTCCCTCGCAT 3773
QY 464 rTyrSerAlaProGlnThrProLeuSerProThrPro 476
Db 3772 CCGATCC-----AAGTCGAATCAATCAAAATCCA 3742

```

RESULT 14
ABLI2636/c
ID ABLI2636 standard; cDNA; 3592 BP.
AC ABLI2636;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32390.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
OS Drosophila melanogaster.
XX
PN WO2001/1042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-0509231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX

PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
DR P-PSDB: ABB68533.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
PS
XX Claim 1; SEQ ID NO 32390; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI1840-ABLI16175), expressed DNA
CC sequences (ABLI1840-ABLI16175) and the encoded proteins
CC (ABBS7737-ABBS7072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp://ipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3592 BP; 1064 A; 781 C; 774 G; 973 T; 0 other:

Alignment Scores:
Pred. No.: 2 656-72 Length: 3592
Score: 828.00 Matches: 184
Best Similarity: 50.66% Conservative: 77
Best Local Similarity: 35.72% Mismatches: 140
Query Match: 26.54% Indels: 114
D: 23 Gaps: 11

US-09-622-964-3 (1-585) x ABLI2636 (1-3592)

```

QY 1 MethrhllethryrThrSerGlnValAlaAsnAlaArg-----LeuGly 15
Db 2904 ATGACTGTGACCTTACTCTCAGTGTGATACCTCGAGCTTCAGTAATTTCTTGGGG 2845
QY 16 SerPheSerArgLeuLeu----- 22
Db 2844 CTGCTGTGACGTTGGTGTACTCAATATATATATATTTTCATTTTAAGCTTTAAAG 2785
QY 23 -----CysTrpArgGlySerIleTyrLysLeuLeuTyrGlyGluPheLeu 37
Db 2784 GAACACCTATTTCAGAGATGCGCGCGATTTTACAGTTAATATGTTGGACCTTAT 2725
QY 38 IlePheLeuLeuCysTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGlnGlu 57
Db 2724 GTATTTTTCGGATTGACTGTTTACTGCGAATTACATATACGTCGTGTTAAGCAAGAG 2665
QY 58 GlnGln-----LeuMet 61
Db 2664 GCGAAAGTAAAGCGTTTATTTGTATTCCTTTTAAATACAAAGCTTTGTACAGGTT 2605
QY 62 PheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPheVal 81
Db 2604 TTGGAGGCATATATGCGATTCGCGAAGAAATAGCTCCCTCATTCCTTATCCGTT 2545
QY 82 LeuGlyPheTyrValThrLeuValValThrArgTyrTrpAsnGlnTyrGlnAsnLeuPro 101
Db 2544 TTGGATTTTTCGTGAGATCTGTATGTAACGATGTTGGACCATATACCACTTCT 2485
QY 102 TrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluSer 121
Db 2484 TGGCCAGATGAATTCATTTTGTATAGCAACCATGATCATGCTCCGATGATAGACCC 2425
QY 122 ArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuLeuArg 141
Db 2424 AGGGTAATGAGCGCAATATCTTAAGATATGTGCGCTGCCAGGTGATATGATTCACC 2365
QY 142 SerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAlaGly 161

```

```
Dh 2364 ATGATATACCGCGGTGTCAAGCGCCGATTCGCCATCACTCAATATTATGAGCTGCC 2305
Oy 162 PheMetThrProAlaGluHisLysGlnLeuGluLysLeu-----SerLeu 176
Dh 2304 TTTCGTGGTGAATAAATGAAAGAAATCATGATGAGCCATGATCAGAGCTTCCCAAGTTAT 2245
Oy 177 ProHisAsnMetPheThrPValProThrPValTrrPheHisAsnLeuSerMetLysAlaTrr 196
Dh 2244 CCAAGAC-----TGGATGCCCATCGCTTTGGCAGCTATGATTTGATGATGAGCGCGG 2191
Oy 197 LeuGlyLysArgLysArgAspProLLeuLeuGlnSerLeuAsnGlnLumetAsnThr 216
Dh 2190 AGGAGATGATAAATCCCGGATGACTACGCTGTTAAGACATCATGATGATGAAACCG 2131
Oy 217 LeuArgThrGlnCysGlyHisLysLeuTyrrAlaTyrAspTrrLLeuSerLLeuProLeuValTyr 236
Dh 2130 TTACGGCGGAATGTGATTTTGTCTCTATTCAGATGATGATGATGATGATGATGATGAT 2071
Oy 237 ThrGlnValValThrValAlaValAlaValSerPhePheLeuThrCysLeuValGlyArgGln 256
Dh 2070 ACCCAAGTGTACACGTCGCGACGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2011
Oy 257 PheLeuAsnProAlaLysAlaTyrProGlyHisGlnLeuAspLeuValValProValPhe 276
Dh 2010 TGGATGAAAGTCATTCGACGAGATGCTACTAGATCAGAGATGCTTCCGATCTCG 1951
Oy 277 ThrPheLeuGlnPhePhePheTyrValGlyTrrPheLysValAlaGlnLLeuLeuAsn 296
Dh 1950 ACCGTCCTCCAGTTTCTTTTACATGAGTGTGCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1891
Oy 297 PropheGlyGlnAspAspAspAspPheGluThr----- 307
Dh 1890 CCTTCGGAGAGATGATGACACTCGAGGTGAGTGTGATTGATTGATTGATTGATTGATTGAT 1831
Oy 307 ----- 307
Dh 1830 TCCCAATGATATACATGTTTCTCTACTTATACATATAGCTTAATGCTTATTATTACTG 1771
Oy 308 -----AsnTrrPLeuValAspArgAsnLeuGlnValSerLeuLeuAlaValAspLumet 325
Dh 1770 CAGCTCACTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1711
Oy 326 HisGlnAspLeuProArgMetGlnProAspMetTyrTrrPAsnLys----- 340
Dh 1710 CACCGAGAGCATCCGGAACCTGTTAAGATCAGATTTGGGAAGAGCTCTCCCTAATGAA 1651
Oy 341 ---ProGluProGlnPro-----ProTyrTrrAlaAlaSerAlaGlnPhe 354
Dh 1650 ATTCCATAGCTGACGCAAAATGCGTCAAAATCCTCCAGACGCTTCCAGCGCTATATG 1591
Oy 355 ArgAlaGlnSerPheMetGlySerThrPheAsnLLeuSerLeuAsn----- 369
Dh 1590 GAATCACAATCAAAAGTAACTTATTTTGTATTTCTCATTAATGATTAAGTAAATATC 1531
Oy 370 -----LysGlnLumetGlnPheGln-----ProAsnGln 379
Dh 1530 CGGATTTTATGCTCCAGAAAGCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1471
Oy 380 GlnAspArgLysLysAlaHisAlaGlyLLeuLeuArgPheLeuGlnLeuGlnSerHis 399
Dh 1470 GAATCAACTTCCGACAAA-----ACTTCTCTGG 1444
Oy 400 AspHisLysProProArgAlaAsnSerArgThrLysLeuLeuTrr 414
Dh 1443 GAAGACCATCTGCGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1399
RESULT 15
ABL12608
ID ABL12608 standard; cDNA; 3785 BP.
AC ABL12608;
XX
DT 26-MAR-2002 (first entry)
```

```
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32306.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-0509231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI, 2001-656860/75.
XX DR P-PSDB; ABB68505.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 32306; 21pp + Sequence Listing; English.
XX PS
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX CC sequences (AB101840-AB116175) and the encoded proteins
XX CC (AB157737-AB172072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pcl_sequences.
XX SQ Sequence 3785 BP; 1064 A; 764 C; 851 G; 1106 T; 0 other.
Alignment Scores:
Pred. No.: 1,46e-69 Length: 3785
Score: 801.00 Matches: 178
Percent Similarity: 53.04% Conservative: 84
Best Local Similarity: 36.03% Mismatches: 136
Query Match: 25.67% Indels: 97
DB: Gaps: 11
US-09-622-964-3 (1-585) x ABL12608 (1-3785)
Oy 1 MetThrLLeuThrTyrTrpSerGlnValAlaAsn---AlaArgLeuGlySerPheSerArg 19
Dh 1001 ATGACTGTCTCTTACACCGGTGAGTGCAACATGACAGCATTTCCGCTCTTCTGGAAG 1060
Oy 20 LeuLeuLeuCys----- 23
Dh 1061 CTTTGTATGAG-GTGAGGCATCTCCGATTCGCCCAAGTTACTAAGATATCGA 1119
Oy 24 -----TrrArgLysSerLLeuTyrrLysLeuTyrrGlyLupheLeuLLeuPhe 39
Dh 1120 TCATTGGCAGATGGGGCCAGTATCTCAAGATATATATGGGTGATCTTCTGCAATTC 1179
Oy 40 LeuLeuCysTyrTrrLLeuArgPheLLeuTyrrArgLeuAlaLeu----- 54
Dh 1180 CTGCTCTGCTCTACTTCATGCGCTGTAATATATGCTATGCTCTCAGGAGCGATGATAAG 1239
Oy 55 -----ThrGlnL 57
Dh 1240 CCGTAAGTATGATATGATATGATGAAAGCTTTTAATTATTATTAACCCCAACCCCTT 1299
```

QY 57 uGlnGlnLeuMetPheGluYulysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIlePr 77
Db 1300 AACCCATAGTCTTTTCGAGCAACATGTACTGATCATATGATGCAACCGATTC 1359
QY 77 oLleSerPheValLeuGluPheTyrValThrLeuValValThrArgTPrPheGlnTyr 97
Db 1360 GCTCTCTTCTGCTACTGGGATTTTATGCGCATATATATGAAACGCGTGGGATCAATA 1419
QY 97 rGluAsnLeuProTPrProAspArgLeuMetSerLeuValSerGlyPheValGluGly 117
Db 1420 TATCACAGTTCCTGGCCGATCCGCGGCTTTATGATGAGCCCTGCTCCGTGGCCA 1479
QY 117 saSPGlnGlnSerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnVal 137
Db 1480 AGATGAACAATGCGTCTGATGAGACGACGATTAATGATAGATATGCTGCTGGCAATTGAC 1539
QY 137 lleuIleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLe 157
Db 1540 TATGGTCTCTGATGATATCGCATTTAAAGCTCCCTTCCAACTACGATCAGTT 1599
QY 157 uValGlnAlaGlyPheMetThrProAlaGluHisLysGlnLeuGluYulysLeuSerLeu 176
Db 1600 AATTGAGTGGGTGCTTAACGCCAACGCAATATTAATGAGCAATGATGTGA 1659
QY 177 -----ProHisAsnMetPheTPrValProTPrValTPrPheAlaAsnLeuSerMetly 194
Db 1660 GTTTCCAAAGCACCCAAAGTATGATGATGCCCTTGTCTGGCGCCGCTATTGTAAACAAG 1719
QY 194 sAlaTPrLeuGluGlyAlaGlyLeuArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMe 214
Db 1720 GGCTGAAAGAAAGGATCGAATTTGGGAGACTTCTCCGTAAGTCCATGATGATGAGT 1779
QY 214 tAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTPrIleSerIleProLe 234
Db 1780 CAATTAATTCGAGTGTGATCATATGCTCATCATGATCAATTCATCAATCTCAGTGCAC 1839
QY 234 uValTyrThr----- 237
Db 1840 GGCTTAACACAGGTAGTTATGACTATTTCAAGCATATATTACGATAGATTATTC 1899
QY 238 -GlnValValThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPh 257
Db 1900 TCAGGTGCTACCCCTGGCGCTACTGCTACTTGTGGCTCGATATTGTGCACCACTG 1959
QY 257 eleuAsnProAlaLysAlaTyrProGlyHisGluLeuAspLeuValValProValPheTh 277
Db 1960 GATCGATCGGATATTACGACTACAAATTAATTTAGTTACTACTTCCGTTGTTCAG 2019
QY 277 rPheLeuGlnPhePhePheTyrValGlyTPrLeuLysValAlaGluLeuIleAsnPr 297
Db 2020 CACTCTGGAGTTTCTTCTTCATGGCGCTTAAGGTGGCCGAGACATTTATTTGTCC 2079
QY 297 oPheGlyGluAspAspAspPheGluThrAsnTPrIleValAspArgAsnLeuGlnVal 317
Db 2080 TTTTGGCATGACGACGATGACTTGAATGAACTGGCTGATGATGCAACTTGCACAT 2139
QY 317 lSerLeuLeuAlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetly 337
Db 2140 GAGTTACTTGAATGATGAAATGACAAATGATCATCCGCAATGGGTGAGGATCAGTA 2199
QY 337 rTPrAsnLysProGluPro-----GlnProTPrTyrThrAlaAlaSer-----AlaGlnPh 354
Db 2200 TTGGGATGAGTGTTCCTCCGCTGAGTTGCGCATATGCCGTGAGTCCGATAGGCCGCAACA 2259
QY 354 eArgAlaGlnSerPhe-----MetGly-----SerTh 363
Db 2260 TCCGAGGATCAACCGCTCATTTGGCATACCAAAATTTGCTCCCGTACCATGACGAA 2319
QY 363 rPheAsnLysSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnGluAspGluGlu 383
Db 2320 AAGCAAGTCAAGTTGGAAAAAGATTTCACAGAGTTC-----GACGATGAGGA 2367

QY 383 uAspAlaHisAlaGlyIleIleGlyArgPheLeuGluGlyLeuGlnSerHisAspHisIlePr 403
Db 2368 TGAATGACAACCCGAGAGTAACTATAGCATTT----- 2398
QY 403 oProArgAlaAsnSerArgThrLysLeuLeuTPrProLys 416
Db 2399 -----GCCCGCGGAGTTCAGTTGAGCAAA 2425

Search completed: July 27, 2003, 11:27:20
Job time : 428.894 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:10:16 ; Search time 2200.84 Seconds

(without alignments)
4054.679 Million cell updates/sec

Title: US-09-622-964-29
Perfect score: 2944
Sequence: 1 MTITYNKRVANARLGSFSL.....EHAESYPRDAGTKPVLYE 551

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xih
-Q/cgn2_1/USPFO.spool/US09622964/rnat 22072003_101119_26026/app_query.fasta_1.2069
-DB-EST -QPM-fastap -SUFIX-rst -MINMATCH-0.1 -LOOPEC-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-DIsum62 -TRANS-human40.cdi -LIST-45
-DOALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFM-pts -NOR-ext HEAPSIDE-500 -MINLEN-0 -MAXLEN-2000000000
-USFR-US09622964.ecgn_1.14188.etrnat.22072003_101119_26026 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -Fgapop-6
-Fgapext-7 -YGAPOP-10 -YGAPEXT-0.5 -DELop-6 -DELEXT-7

Database : EST.*
1: em_estda.*
2: em_esthum.*
3: em_estin.*
4: em_estinu.*
5: em_estrov.*
6: em_estrpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_hic3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2621	89.0	1699	11 AK006549	AK006549 Mus muscu
2	1520.5	51.6	2453	11 BC015220	BC015220 Homo sapi
3	1165	39.6	963	14 B0879880	B0879880 AGENCOURT
4	786	26.7	666	13 B1480798	B1480798 H2RPE-043
5	786	26.7	735	13 B1756228	B1756228 603024265
6	779	26.5	793	13 BG961794	BG961794 602826439
7	736.5	25.0	654	10 BE679896	BE679896 d558a11.y
8	711	24.2	508	10 BE236901	BE236901 145906 MA
9	701	23.8	526	14 BM685501	BM685501 UI-E-C10-
10	670	22.8	522	14 BM707649	BM707649 UI-E-C11-
11	664	22.6	613	9 AL654606	AL654606 AL654606
12	647	22.0	566	14 BM718338	BM718338 UI-E-E01-
13	634	21.5	628	13 B069061	B069061 B069061
14	634	21.5	722	9 AL627604	AL627604 AL627604
15	628	21.3	566	13 B1443895	B1443895 GA196B11.
16	625	21.2	732	13 BG965806	BG965806 602829577
17	592	20.1	440	12 BG084076	BG084076 H3095G07-
18	580	19.7	657	14 BQ397157	BQ397157 NISC-ng26
19	570.5	19.4	672	13 BM576476	BM576476 170006871
20	561	19.1	577	10 BE669309	BE669309 dc57f06.y
21	557	18.9	1067	13 BM562042	BM562042 AGENCOURT
22	554	18.8	455	12 BG013943	BG013943 IL5-GN024
23	553.5	18.6	821	12 BE793366	BE793366 601588385
24	548	18.6	966	12 BE583222	BE583222 602094224
25	539	18.3	635	14 BQ396082	BQ396082 NISC-ng18
26	539	18.3	1111	17 BQ396082	BQ396082 NISC-ng18
27	529	18.0	649	10 BE189780	BE189780 db60a04.y
28	511	17.4	1123	14 BQ276342	BQ276342 AGENCOURT
29	510	17.3	605	14 BQ391968	BQ391968 NISC-ng21
30	508	17.3	736	12 BG071195	BG071195 H3095G07-
31	503.5	17.1	550	13 BG663329	BG663329 602827751
32	496	16.8	514	13 B1343182	B1343182 371242 MA
33	493.5	16.8	504	9 AA205892	AA205892 zq47a06.r
34	478	16.2	508	12 BG892937	BG892937 daa91d01.
35	466.5	15.8	1064	17 CNS02AOC	CNS02AOC Terradon
36	462	15.7	755	13 BG923094	BG923094 Terradon
37	460	15.6	621	9 AL672519	AL672519 AL672519
38	457.5	15.5	578	9 AU040308	AU040308 60282917
39	457	15.5	271	9 AA497726	AA497726 v172b03.r
40	455	15.5	797	10 AV398472	AV398472 AV398472
41	450	15.3	531	14 BM685396	BM685396 UI-E-C10-
42	447	15.3	930	14 BQ436824	BQ436824 AGENCOURT
43	447	15.2	525	14 BQ345562	BQ345562 PM3-NY031
44	444	15.1	614	9 AL782195	AL782195 AL782195
45	434	14.7	470	9 AA573517	AA573517 nk99g10.s

ALIGNMENTS

RESULT 1
LOCUS AK006549
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700030H21:vtelliform macular dystrophy 2 homolog (human), full insert sequence.
ACCESSION AK006549
VERSION AK006549.1 GI:12839710
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone:1700030H21.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

Db 361 CCACA -AACACATTCTGGTGGCTTGGGTGGTGGTTGGCCAACTGTCAATGACAGGCCAT 4119

Qy 197 LeuGIyGIyRtIleArGAspThrValIleuLeuGInSerIleMetAsnGIyValCysThr 216

Db 420 CTTGGAGGCTCGAATCCGGGACACCGCTCCGCTCCAGACCTGATGAATGAGGTGTACT 479

Qy 217 LeuArGtGInCysGlyGInLeuIleuYrAlaYrAspTrpIleSerIleProIleValTyr 236

Db 480 TTGCGATCACTAGGTGGACAGCTGTATCTACAGACTGGATAGATATCCATTGGTGTAC 539

Qy 237 ThrGIyValIleThrValAlaValYrSerPhePheLeuAlaCysLeuIleGIyArGIn 256

Db 540 AACACAGTGTGACAGTGGACAGTATACAGCTTTTCTTGAGTCTTGACGGGAGCAG 599

Qy 257 PheLeuAsnProAsnIleYrAspTyrProGIyHisGIuMetAsnLeuValIleProAlaPhe 276

Db 600 TTTCGAAACCCAAACAAAGAGCTATCCAGAGCCATAGATGATCGGTGGCTGTCTTC 659

Qy 277 ThrIleLeuGInPheLeuPheTyrMetGIyTrpLeuYsValIleGIuGInLeuIleAsn 296

Db 660 ACAATCCGCAATCTTATTCTACATGGCTGGCTGAGGGTGGCAACAGCTCAACAC 719

Qy 297 ProPheGIyGIuAspAspAspPheGluThrAsnTrpIleIleAspArgAsnLeuGIn 316

Db 720 CCCTTCGGGGAGAGCAGATGATGATTTGAGACTACCTGATCATTCAGCAAACTGCAG 779

Qy 317 ValSerLeuLeuSerValIaSPGIyMetHisGIuAsnLeuProIleMetGIuArGAspMet 336

Db 780 GTGTCCCTGTGTCCGTGGATGGGATGCCACAGAACCTTCCTCCATGGAACTGACATG 839

Qy 337 TyrTrpAsnGIuAlaAlaProGIuProProTyrThrAlaAlaSerAlaArgSerArgArg 356

Db 840 TACTGGAAACAGAGGAGGCTCAGCCGCTCAGCCCTCAGACAGCTGTTCGCCAGTTCCGCCG 899

Qy 357 HisSerPheMetGIySerThrPheAsnIleSerIleuYsGIuAsnLeuGInLeuTrp 376

Db 900 CATCTCTTACGTGGCTCCACACTTCAACATCACTCAAAAGAAAGAAAGATTGTGACTTTGG 959

Qy 377 SerIleGIuGInIuAlaAspThrAspIleYsGIuSerGIyTyrSerSerThrIleGIyCys 396

Db 960 TCAAAAGAGAGGCTGACACGATTAAGAAAGAGAGTGGCTTACAGCACCATAGGCTGC 1019

Qy 397 PheLeuGIyLeuGInProIleYsAsnTrpHisLeuProLeuYsAsnLeuYsThrIleu 416

Db 1020 TTCTTGTAGACTGCAACCCAAAACATCACTCTCCCTTGAAGAAAGCTTAAAGACCAACAT 1079

Qy 417 IleuYsSerIleYsAsnProLeuLeuGIyGInCysIleYsAspAlaAsnGIuYsAsnGIu 436

Db 1080 TTGTGTCTTAACAAACCCCTCTCGAAGGCCAGGTAAAGATGGCAACAGAAAACCAAG 1139

Qy 437 YsAspValIleTrpIlePheIleYsGIuLeuAspPheLeuIleCysValProArgPheYsArg 456

Db 1140 AAAAGATGCTGGAATTTAAAGGCTGTGGACTCTTGAATAGTCTTCCAAAGGTTTAAAGAG 1199

Qy 457 ArgGIySerHisCysGIyProGIuAlaProSerSerHisProthrGIuGInSerAlaPro 476

Db 1200 AAGGCTCCCATTTGGGCCACAGGACCCAGCAGCACCACTTACTAGAGATCACACACC 1259

Qy 477 SerSerSerAspTrpHisGIyAspGIyProSerThrAspTyrGIuIleCysHisMetIleYs 496

Db 1260 TCCAGTTACAGACAGAGTGGGCTTCCACAGATTACCAAGAAATCTGTCACTGATAA 1319

Qy 497 YsIleYsThrValGIuPheAsnLeuAsnIleProGIuSerProThrGIuHisLeuGIn 516

Db 1320 AAGAAACTGTGGAGTTTAACTTGAACATTTCCAGAGAGCCCAACAGAACATCTTCACAG 1379

Qy 517 ArgArgLeuAspGIuMetSerThrAsnIleGIuAlaLeuMetIleYsGIuHisAlaGIuSer 536

Db 1380 CGCGCTTTGGACACAGATGCTCAACCAATTAACAGGCTCTTAATAAGAGACATGCAGAGTCC 1439

Qy 537 TyrProTyrArgAspGIuIleGIyThrIleYsProValIleuTyrGIu 551

DB	1440	TATCCCTACAGGATGAAGCTGGCACCAACCTGTCTATGAG	1484
RESULT 2			
BC015220			
LOCUS	BC015220	2453 bp	mRNA
DEFINITION	Homo sapiens, clone IMAGE:3877806,		linear
ACCESSION	BC015220		HTC 25-JUL-2002
VERSION	BC015220.1	GI:21955361	
KEYWORDS	HTC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2453)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
	Submitted (01-OCT-2001) National Institutes of Health		

QY 112 serPheVal¹GIu²GIy³Leu⁴Asp⁵Leu⁶GIu⁷GIu⁸Val⁹Arg¹⁰Leu¹¹Arg¹²Leu¹³Arg¹⁴Thr¹⁵Leu¹⁶Ile¹⁷Arg¹⁸Val¹⁹ 131
 Db 738 GCGTTCGTGGAAAGGCGAAGAGACAGACAGCGCGGTGGCTGGCGGCGACAGCTCATCGCTAC 797
 QY 132 AlaIleLeuGIuGIuValLeuIleLeuArgSerIleSerThrSerValIleValIleArgPhe 151
 Db 798 GCCAACCTGGGCAACGTGCTCATCTCTGGCGAGCGCTCAAGACCGGACAGTCTCAAGCGCTTC 857
 QY 152 ProThrLeuHisIleLeuValLeuAlaGlyPheMetThrHisGIuHisIleValGIuLeu 171
 Db 858 CCCAGCGCCGACACCTGGTCCAGACAGCGCTTTATGACTCCGGCAGAACACAGCACTTG 917
 QY 172 GIuHisLeuGIyLeuProHisAspThrPheThrValProThrValIleThrPheAlaValLeu 191
 Db 918 GAGAAAGTACGACCTTACACACACACATCTTCTGGGTGGCGGGGTGGTGGTGGCAACTTG 977
 QY 192 SerMetIysAlaValLeuGIyGIyArgIleArgAspThrValLeuGIuGIuSerLeuMet 211
 Db 978 TCATATGAAGGGGTGGCTTGGAGGTCGAATCCGGGACCCATTCCTGCTCCAGACCTTGCTG 103
 QY 212 AsnGIuValCysThrIleArgThrGlnGlyGlnLeuValAlaValAspTrpIleSer 231
 Db 1038 AACGAGTGAACACCTTGCTACTCAAGTGGACACCTGATCTCAACAGCTGATTAAT 109
 QY 232 IleProLeuValIleThrGlnValIleValIleValIleValIleValIleValIleValIle 251
 Db 1098 ATCCCACTGGTGTATACACAGGTGGTACGTGGCGGCTGACAGCTTCTTCTGACTGT 1157
 QY 252 LeuIleGIyArgGlnPheLeuAsnProValAspTrpProGIyHisGIuMetAspLeu 271
 Db 1158 CTAGTGTGGGGGCGAGTTTCTGAACCCAGGACGAGGCTTACCTTGGCATGAGCTGGACCTC 1217
 QY 272 ValValProValPheThrIleLeuGIuPheLeuPheThrMetGlyThrLeuValAla 291
 Db 1218 GTTGTCCCGGTTCACGTTCCGACGTTCTTCTATGTTGGCTGGCTGAAG---- 1271
 QY 292 GIuGlnLeuIleAsnProPheGIyGIuAspAspAspPheGIuThrAsnTrpIleIle 311
 Db 1271 ----- 1271
 QY 312 AspArgAsnLeuGIuValSerLeuSerValAspGIuMetHisGIuAsnLeuProPro 331
 Db 1272 -----GTGTCCTGTGGCTGTGGTGTGGATGTACACAGACCTCGCTCGG 1316
 QY 332 MetGIuArgAspMetCysTrpAsnGIuAlaAlaProGlnProProTyrThrAlaIleSer 351
 Db 1317 ATGGAGCCGACATGTACTGTGAATATGAACCCAGACACAGCCCTTACACAGCTGCTTC 1376
 QY 352 AlaArgSerArgArgHisSerPheMetCysSerThrPheAsnIleSerLeuValGIu 371
 Db 1377 GCCCAGTTCCGTCGAGCTCTTTATGGGCTCCACCTTCAACATCACACCTGACAAAGAG 1436
 QY 372 AspLeuGIuLeuTrpSerGIuGIuAlaAspThrAspValGIyGIuSerGIySer 391
 Db 1437 GAGATGGAGTTCCAGCCCAATCAGAG-----GACGAGAGATGCTCAAGCT 1484
 QY 392 SerThrIleGIyCysPheLeuGIyLeuGIuProLysAsnTrpHisLeuProLeuLysAsp 411
 Db 1485 GGCAATCATGGCCGCTTCTAGGCTCCAGTCCCATATCACATCTCTCCAGGAGCA 1544
 QY 412 LeuIleThrIleLeuCysSerIleAsnProLeuLeu-----GIuGIuGIuValGIy 429
 Db 1545 TCAAGGACCAAACTACTGTGGCCCAAGAAGGAATCCCTTCCACAGAGGCTCGCCCAAA 1604
 QY 430 Asp-----AlaAsnGIuLysAsnGIuLysAsp-----ValTrpIlePhe 442
 Db 1605 AACCAACAAGCAGCAACAACAACGTTTGGGGCCAGAGAACAAACAAGCCTGGAGGCTT 1664
 QY 443 LysGIyLeuAspPheLeuLysCysValProArgPheValArgIleValSerHisCysGIy 462
 Db 1665 AAGGCTGTGGACGCTTCAAGTGTGGCCCACTGTATCAAGGCGCAAGGCTACTACATGTGCT 1724

```

QY      463 ProGlnAlaProSerSer-----HisProThrGluInSerAlaProSer 477
      ||||| ||| |||
Db      1725 CCACGACACCCCTCAGCCCACTCCCATGTTCTTCCCTCCTAGAACCATCAGCGCGTCA 1788
QY      478 Ser-----SerAspThrGly----- 482
      ||| |||||
Db      1785 AAGCTTCACAGTGTACACAGCATACACACCAAAGACAAAGCTTAAAGACTGTGATTCT 1844
QY      483 -----AspGlyProSerThrAspTyrGln 490
      |||||
Db      1845 GGGGCGCAAGAAAGTTTGAAATTGCTGTCCAGAGCAGCATGGGCGCTTGATGAGCACCCA 1904
QY      491 GlnIleCysHisMetLysLysLysThrValGluPheAsnLeu--AsnIleProGluSer 509
      ||||| ::::: ||||| |||||
Db      1905 GAAAGTATCTCAAGTCAGGAGGAGAAACTGTGAGATTAACTCCAGACGAAATCCACAGATC 1964
QY      510 ProThrGlnHisLeuGlnIleArgArgLeuAspGlnMetSerThrAsnIleGlnAlaLeu 529
      ||| ||||| ::::: ||||| |||||
Db      1965 CCGGAAATACACCTCAAGAAG--CCTTTGGACATATCACCACCAACCATATCACACTACA 2021
QY      530 MetLysGluHisAlaGluSerTyr-----ProTyrArgAspGluAlaGlyThrLys 546
      ::|||::||| ::||| ||||| |||||
Db      2022 CTCAAAGATCACATGATCTTATTGGCGCTTGGAAAAACAGGATGAAGCATATTCCTAA 2081
QY      547 Pro 547
      |||
Db      2082 CCT 2084

```

FEATURES

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/clone_1lb="Lupsk1_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/notes="Vector: PCWV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dt priming.
directionally cloned using the following adaptors:
5'-TCGACCCACGGCGTCGG-3' and
5'-GACATGTTTGATGATCGAGCGCCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupsk1, M.D./Ph.D. (Baylor

```



```

FEATURES
source      Location/Qualifiers
1..793

```

Alignment Scores:	
Pred. No.:	1.72e-79
Score:	779.00
Percent Similarity:	75.48%
Best Local Similarity:	60.15%
Query Match:	26.46%
DB:	13
	13
Length:	79
Matches:	157
Conservative:	60
Mismatches:	43
Gaps:	3
Gapels:	3

00-00000-904-29 (1-551) X BG961794 (1-793)

```

Oy 77 ProIleSerPheValLeuGlyPheTyrValThiLeuValIleSerArgTyrPrpSerGln 96
Db 13 CCGGTCTCTTTGCTACTGTTGGCTTTACGTACTGTCGGTGGCATCCCTGTGTGAACAG 72
Oy 97 TyrGIaAsIleuProTyrProAspArgIleuMetIleGlnValSerSerPheValIGlnGly 11
Db 73 TACCTATGCATGCTCTCTGCGGACGACCATCATGTGCATAGTGGGTGGACCCGGCATGGG 13
Oy 117 LysAspGluIGluGlyArgLeuLeuAspArgThiLeuIleArgTyrAlaIleuGln 13
Db 133 CGAGAGCATCGAGGGCGGCTCTACCGGGGCAAGCTCATGGCTACGACAGGGCTCTCGGG 19
Oy 137 ValLeuIleLeuAspSerIleSerThSerValTyrTyrAspArgPheProThiLeuHis 15
Db 193 GTGGCATCTCTCGTTGTTGTAAGCAGACAGCTCTCAAGGGTCTCCCACTTAACACAC 25
Oy 157 LeuValIleuAlaGlyPheMetThrHisGlyGlnIleLysGlnLeuGlnLysLeuGlyLeu 17
Db 253 GTGGCGAGAGCTGGATTTATGACCCCGAAGAGACGCCAAGAGTGGAGAACTGTGAATCG 31
Oy 177 ProHisAsnThrPheTyrValProTyrValThrPheAlaAsnLeuSerMetLysAlaTyr 19
Db 313 TCTACCAACAATACTGGTGGCTCGCTGATGGTTCTTCAGGCTGGCGACCGACGGCGGG 37
Oy 197 LeuGlyLysArgIleArgAspThrValLeuGlnGlnSerLeuMetAsnGlnValCysThr 21
Db 373 CGAGAGGGGGGCATCCGCGACACAGTGGCCATAAGTTGCTACTGAGACACTTAATGTG 43
Oy 217 LeuArgThrGlnCysGlyGlnLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyr 23
Db 433 TTTCGGAGCAAACTGGGATGCTGTTCCTACCTAGCAGACTGATTAACATACCCCTGCTAC 49
Oy 237 ThrGlnValValThrValAlaValLysSerPhePheLeuAlaCysLeuIleGlyArgGln 25
Db 493 ACCCAGGTATGCTACTGCCACTGTACAGCTACTTGTGGCTCCCTCATCGGCGCGAC 55
Oy 257 PheLeuAsnProHisLysPyrTyrProGlyHisGluMetAspLeuVal-ProValPhe 27

```

Db	553	TTCTGAGACCTTCACAGGGCTACAAAGCACACCCCTGGACCTTATGCGTACCCCATCTT	612
QY	276	ethrlleleuglnlPheleupherymetglYtPleuLyValAlaglnuInleuIleas	286
Db	613	CACCTGTGCTGCAAGTTCTTCTTCAAGCTGGCTGGCTTAAAGTAGCAGACGACTTATTAA	672
QY	296	nPropheglYgluaspaspaspPheglutIraanrTPIlelleleasPArgasnleuI	316
Db	673	-CCCTTTGGAGAGGAGCAGAGACGACTTTCAGACCAATT-CTTATGACCGAAGAAATCCCA	730
QY	316	nValserleuLeuSerValaspIYmerHlglnInleuPProPromeclIuArgaspme	336
Db	731	GGTGTGATGATGTAGCTGTAAAGATGAAATGATGATATACCACTGGCCACAGCTGGCAAAAGATCT	790
QY	336	t 336	
Db	791	A 791	

RESULT 7	LOCUS	DEFINITION	EST 31-MAY-2002
BE679896	BE679896	BE679896	
		654 bp	linear
		df58a11.y1	Xenopus laevis
		unfertilized egg	CDNA library
		laevis	CDNA library
		SHR.VM2_HUMAN	O76090
		BESTROPHIN 1	mRNA sequence.
		BE679896.1	GI:10063113
		EST	
		African clawed frog.	
		Xenopus laevis	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 654)	Clifton, S., Johnson, S. L., Blumberg, B., Song, J., Hillier, L., Page, M., J. J., Wylie, T., Underwood, K., Theisling, B., Bowers, Y., Pers, B., Gibbons, M., Harvey, N., Rittler, E., Jackson, Y., McCann, R., Waterson, R. and Wilson, R.	Wasnu Xenopus EST project, 1999	Unpublished (1999)	Other ESTs: df58a11.x1

FEATURES
source

Email: est@watson.wustl.edu
Library constructed by Bruce Blumberg
Library normalized by Jitwan Song
DNA Sequencing by: Washington University Genome Sequencing Center
This clone is available royalty-free through LLNLT, contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Glc3c
High quality sequence stop: 418.
Location/Qualifiers
1. 654
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XENOPUS.-SOURCE.ID:xlineg003a21"
/clone_lib="Xenopus laevis unfertilized egg cDNA library"
/tissue.type="unfertilized egg"
/lab.host="Top-10 F."
/note="Vector: pBluescript SK-. Site_1: EcoRI; Site_2:
XhoI. This library has been used successfully to clone a
number of full-length cDNAs ranging in size from 1.4 to
4.5 kb. There are less than 0.5% clones with multiple
inserts. Since each cDNA has an EcoRI site at its 5' end
and an XhoI site at the 3' end, these clones can be easily
identified. One should be suspicious of any clone which
gives 3 or more bands in a XhoI-XhoI double digest AND
has an internal XhoI site. We usually do not further
characterize any such clones unless the cDNA is known to
give multiple bands in an EcoRI-XhoI digest. Microplate

BASE COUNT	107 a	145 c	148 g	126 t
ORIGIN				
Alignment Scores:				
Pred. NO.:	1.04e-70	Length:	526	
Score:	701.00	Matches:	134	
Percent Similarity:	83.62%	Conservative:	14	
Best Local Similarity:	75.71%	Mismatches:	24	
Query Match:	23.81%	Indels:	5	
DB:	14	Gaps:	1	
US-09-622-964-29 (1-551) x BM685501 (1-526)				
QY	222	GLYGLNLEUTRYALATYRSPRRPPILESERLIEPRLLEU-VALYTRHGINVALVATH	241	
DB	9	GGACACCGTATGCTCCCTACGACTGGATTGATCCACAGGGGATACACAGCTGGTAC	68	
QY	241	YVALAVALLYTSERPHEPELEALACYSLENIIEGLYARGINPHLEUASPROAS	261	
DB	69	TGTGGCGGTGACACACTTCTCCGACTGTCGTGTGGGGGAGATTCTGAACCCAGC	126	
QY	261	NLYSAPRYPROGLYHISGLIUMETASPLEUAVAIYPROVALPHERHRIELEUNPH	281	
DB	129	CAAGCCATACCTGGCCATGAGCGAGCACTCGTGGCCGCTTACAGTTCGCGAGTT	188	
QY	281	ELEUPHETMETGLYTRIPLEULYSVALAIALAGLNLLEULEASNPOROBEGLYLUAS	301	
DB	189	CTTCTTCTATGTTGGTGGCTGGAAGGTGGCAGAGCAAGCTCATCAACCCCTTGGAGAGA	248	
QY	301	PASAPSPAPHELUHTRASNTPILEILEASPARVASLEUGLIVALSERLEUENSE	321	
DB	249	TGATGATGATTTGAGACCAACATGATGTCGACAGAAATTGACAGTGCTCCCTGTGGC	308	
QY	321	YVALASPGLYMETHISGLINLEUPROKMETGLIAYGAPMETTYTRTPASGLUAL	341	
DB	309	TGTGGATGAGATCACCAGAGACTGCTCGTGGATGGAGCCGGAACATGTACTGGAAATAGCC	368	
QY	341	AALAPROGLINPROPTYTRTHALALASERLAAIYGSERATGATGHIISERPHMETGL	361	
DB	369	CGAGCCACAGCCCCCTACACAGCTGCTCCGCCAGTTCGCTGACGCTCTTATGGG	428	
QY	361	YSETRHPHEANILSERLEULYGLIANSPLEUNLEUTRSPRLYSGIUGLUAL	381	
DB	429	CTCCACCTTCAACATCACCTGTAACAAAGAGAGATGAGTTCAGCCCAATCAGGAG	486	
QY	381	AASPHTRASPLYSGLIUSERGLYTYRSESTHRIIEGLCYSPHE	397	
DB	487	-----GACGAGGAGATGCTCAACGCTGGCATCATATGGCCGCTTC	525	

RESULT 10
 BM707649
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT

BM707649 522 bp mRNA linear EST 28-FEB-2002
 UI-E-C11-af5-o-15-0-01.r1 UI-E-C11 Homo sapiens cDNA clone
 UI-E-C11-af5-o-15-0-01.5', mRNA sequence.
 BM707649
 EST
 BM707649.1 GI:19020907
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 522)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 Contact: Soares, MB

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

QY 146 SerValTyrLeuArgPheProThrLeuHisLeuValLeuAlaGlyPheMetPheHis 165
 Db 249 GGTGCTTCAAAAGATTCCTACATAGACCATGTTGTAGAAAGCTGGGTTATGACTCCG 308
 QY 166 GtlyGlnHisGlnLeuGlnLeuGlnLeuProHisAsnThrPheTrpValProTrp 185
 Db 309 CTGGAGAGAAAGAGTTGAGACCTTACATGTCATCTATACCAAGTACTGGTCCGTC 368
 QY 186 ValTrpPheAlaAsnLeuSerMetLeuAlaTyrLeuGlnGlyArgGlyLeuAspThrVal 205
 Db 369 GTGGGTCTGTACCTGGCTCCGACGGCGGAGCGAGGAGATCCAGATCCAGATCAT 428
 QY 206 LeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgThrGlnCysGlnLeuTyr 225
 Db 429 TCCTTAAATGCGATGAGAGACCTGACACCTTCGCGGGAAGTCCGCGCATGCTTC 488
 QY 226 AlaTrpAspTrpPheSerIleProLeuValTyrThrGlnValAlaValAlaValTyr 245
 Db 489 CACTATGACTGATGATCAGCGCTCCGCTGTTATACACAGTTGTGCAATTCGGGTAC 548
 QY 246 SerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnProAsnLysAspTyrPro 265
 Db 549 AGCTTCTCTGACCTGCTGATGATGAGCCAGCTTCCTGATCCGACGCGGATATCC 608
 QY 266 Gly 266
 Db 609 GGT 611
 RESULT 12
 LOCUS BM718338 566 bp mRNA linear EST 01-MAR-2002
 DEFINITION UI-E-EO1-8ja-j-23-0-UI-r1 UI-E-EO1 Homo sapiens cDNA clone
 ACCESSION BM718338
 VERSION BM718338.1 GI:19036537
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (Bases 1 to 566)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Tissue procurement: Dr. Gregg Hageman
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 FEATURES
 source Location/Qualifiers
 1..566
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-EO1-8ja-j-23-0-UI"
 /clone_id="UI-E-EO1"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pMT3-Pac (Pharmacia) with a
 modified polylinker; Site:1; Ecor I; Site:2; Not I;
 UI-E-EO1 is a normalized cDNA library containing the
 following tissue(s): fetal eye. The library was

constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an Ecor I
 adaptor, digested with Not I, and cloned directionally
 into pMT3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (GT)₁₈ tail. The sequence tag for this library is
 CCGTATACC. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI).
 BASE COUNT 107 a 182 c 145 g 132 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,226-64 Length: 566
 Score: 647.00 Matches: 132
 Percent Similarity: 55.98% Conservative: 13
 Best Local Similarity: 50.97% Mismatches: 14
 Query Match: 21.98% Indels: 100
 Gaps: 1
 US-09-622-964-29 (1-551) x BM718338 (1-566)
 QY 1 MethrIleThrTyrThrAsnLysValAlaAsnAlaArgLeuGlySerPheSerSerLeu 20
 Db 88 ATGACCAATCACTTACACACAGCAAGTGGCTTAATGCGCTTCTCCGCTCCGCTG 147
 QY 21 LeuLeuGlySerTrpArgGlySerIleTyrLysLeuLeuTyrGlyLupheLeuValPheile 40
 Db 148 CTCGTGCGCGGGGCGGACATCTACAAAGCTCTATATGCGAGGTCTCTAATCTTCG 207
 QY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGlnLeu 60
 Db 208 CTCGTCTACTACATCATCCGCTTATTTATAGCGTGGCCCTACGAGAAACAGACGCG 267
 QY 61 LeuPheGlnLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
 Db 268 ATGTTTGAGAAACAGTACTCTATATGCGACACTATACACACCTATCCCATTTCTTC 327
 QY 81 ValLeuGlyPheTyrValThrLeuValAlaSerArgTrpPheSerGlnTyrGlnAsnLeu 100
 Db 328 GTGCTGGCGCTTCTACGTGACGCTGCTGACCCCTGTGGAAACCATAGACGACCTG 387
 QY 101 ProTrpProAspArgLeuMetIleGlnValSerSerPheValGlnGlyLysAspGlnGlu 120
 Db 388 CGGTGGCCGACCGGCTCATGAGCTGTGGGCTTCGTCGAAAGCAAGGACGAGCA 447
 QY 121 GlyArgLeuLeuArgArgThrLeuIleArgGlyAlaIleLeuGlnGlyValLeuIleLeu 140
 Db 448 GGCGGCGCTGCGGCGGCGACGCTCATCCGCTACCGC----- 483
 QY 141 ArgSerIleSerThrSerValTyrLysArgPheProThrLeuHisLeuValLeuAla 160
 Db 483 ----- 483
 QY 161 GlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLeuProHisAsnThr 180
 Db 483 ----- 483
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrLeuGlyGlyArg 200
 Db 483 ----- 483
 QY 201 IleArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgThrGln 220
 Db 483 ----- 483
 QY 221 CysGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValVal 240
 Db 484 -----AACCTGAGTATACACAGGTGCTG 507

OY 241 ThrValAlaValTyrSerPheLeuAlaCysLeuIleGlyTyrGlnPheLeuAsn 259
 DB 508 ACTGTGGCGGTGACACTTCTTCCCTGACTTGTCTACTGTGGCGGAGTTTGTGAAC 564
 RESULT 13
 LOCUS BJ069061
 DEFINITION BJ069061 628 bp mRNA linear EST 11-DEC-2001
 laevis cDNA clone X1051X03 5', mRNA sequence.
 ACCESSION BJ069061
 VERSION BJ069061.1 GI:17494758
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 628)
 Kitayama, A., Terashima, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara
 Y.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadasi Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 FEATURES
 SOURCE
 Location/Qualifiers
 1..628
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="X1051X03"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"
 /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNA
 were oligo-dT primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is substracted
 and was constructed by N. Garrett and A.M. Zorn,
 (Wellcome/CRC Institute)."
 BASE COUNT 146 a 141 c 161 g 180 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8 49e-63 Length: 628
 Score: 634.00 Matches: 117
 Percent Similarity: 76.41% Conservative: 32
 Best Local Similarity: 60.00% Mismatches: 46
 Query Match: 21.54% Indels: 0
 DB: 13 Gaps: 0
 US-09-622-964-29 (1-551) x BJ069061 (1-628)
 OY 1 MetThrIleThrTyrThrAsnysValAlaAsnAlaArgLeuGlySerPheSerIleu 20
 DB 32 ATACAGGTACACCTACACGACGACGAGCAATGCCGTTTGGGGCTTTATTAACGTG 91
 OY 21 LeuLeuCysTrpArgGlySerIleTyrIysLeuLeuTyrGlyLupheLeuValPheIle 40
 DB 92 TTGCTCTCTATGAGGGGGAGTATATACAAATGCTGTACAGAGATTCCTGGCTTTTAA 151
 OY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValIleSerSerAspGlnIleu 60
 DB 152 ATCAGTACTCTGGGCTCTGACATATACAGCTTTTGGTCAATGATGACAGAGACTT 211
 OY 61 LeuPheGluTysLeuAlaLeuTyrIrycysAspSerTyrIleGlnLeuIleProIleSerPhe 80
 DB 212 TACTTTGAGAAAGTGGCCATTACTGCAATACATGCAAACTCATCCTGCTGTGCTTT 271
 OY 81 ValLeuGlyPheTyrValThrLeuValAlaSerArgTyrTrpSerGlnTyrGlnAsnLeu 100

DB 272 GTGTGGGTGTTTATGTACACCTGTGTAAATCGTGTGTAACCACTACTCTGTGTATG 331
 OY 101 ProTTPProAspArgLeuMetIleGlnValSerSerPheIleGluGlyLysAspGlnLu 120
 DB 332 CCTTCCAGACACCGGTGATGTGTGCTGTCCCGTACTGTACTGATGGTGTGATGAAACG 391
 OY 121 GlyArgLeuLeuArgArgThrLeuIleArgTyrAlaIleLeuGlyGlnValLeuIleLeu 140
 DB 392 GGAGCTCTCTACCGCGGACACGTATAGCTACGACGCTTACAGACCTATTAATCCTG 451
 OY 141 ArgSerIleSerThrSerValTyrTyrArgPheProThrIleuThsIleValLeuAla 160
 DB 452 CGTTCGCTTACTACCGGTGCTTCAAAAAGATTCTTCAATACACCATGTGTGAGAGCT 511
 OY 161 GlyPheMetThrIleGlyGlnLysGlnLeuGlnIleLeuGlyLeuPheProHisAsnThr 180
 DB 512 GGGTTCATGACTCCCTGAG 571
 OY 181 PheTrpValProTyrValTyrPheAlaAsnLeuSerMetLysAla 195
 DB 572 TACTGGGTTCGTCGCTGTGTGTCTGTACTGTGACACAGCA 616
 RESULT 14
 LOCUS AL627604
 DEFINITION AL627604 722 bp mRNA linear EST 02-NOV-2001
 mRNA sequence.
 ACCESSION AL627604
 VERSION AL627604.1 GI:16597087
 KEYWORDS EST.
 SOURCE western clawed frog.
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 722)
 Huckie, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (10_2001)
 Unpublished (2001)
 Contact: Huckie E
 Sanger Centre
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TGA8025m15.sp6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 FEATURES
 source
 Location/Qualifiers
 1..722
 /organism="Silurana tropicalis"
 /db_xref="taxon:8364"
 /clone="TGA8025m15"
 /clone_lib="XGC-gastrula"
 /dev_stage="gastrula (stages 10.5-13 mixed)"
 /lab_host="Escherichia coli XL1-blue"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dT primed from 5' end of poly A+ RNA from stages
 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
 into pCS107 with EcoRI at the 5' end and NotI at the 3'
 end."
 BASE COUNT 156 a 177 c 190 g 196 t 3 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1 06e-62 Length: 722
 Score: 634.00 Matches: 125
 Percent Similarity: 71.43% Conservative: 35
 Best Local Similarity: 55.80% Mismatches: 64
 Query Match: 21.54% Indels: 1
 DB: 9 Gaps: 0

Sequence 11070, *P*[illegible]

```

RESULT 1
US-10-198-846-11070
Sequence 11070, Application US/10198846
Publication NO. US20030099974A1
GENERAL INFORMATION:
APPLICANT: lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Stelmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

```

LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
 OTHER INFORMATION: n - A,T,C or G
 US-10-198-846-11070

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4,29e-79	1292	761.50	168	31	51	1	
Percent Similarity:		63.708					
Best Local Similarity:		52.678					
Query Match:		25.878					

US-09-622-964-29 (1-551) x US-10-198-846-11070 (1-1292)

```

QY 82 LeuGlyPheTyrValThrIleuValSerArgTTPSPSerGlnTyrGlnAsnLeuPro 101
Db 317 ATAGGGTTTATGTTACTCTGCTAGTGAACCGATGCGAACCAGTCTGTAATTTGCC 376
QY 102 TTPProSPArgLeuMetIleGlnValSerSerPheValGlnGlyLysAspGlnGly 121
Db 377 TGGCCAGACAGCGCTAATGTCTCTCATCTCTAGCACTGTTACGGAAGCAGCAGCGG 436
QY 122 ArgLeuArgArgThrIleuValLeuArgTyrAlaIleuGlnValLeuArg 141
Db 437 CCGCTGCTTAGAAGACCGCTAGCGCTACGTCATCATCTCACCTCCCTCATCTTCCG 496
QY 142 SerIleSerThrSerValTyrLysArgPheProThrIleuHisIleuValLeuAlaGly 161
Db 497 TCGGAGACACTGCTGTGTACAAAGATTTCCCAATGACCCAGCGGTGAAGCGAGT 556
QY 162 PheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLeuProHisAsnThrPhe 181
Db 557 TTATGACACAGATGAAGAAATTTATTCACCCACTCAAGTCTCATCTGAATAT 616
QY 182 TyrValProTyrValTTPPheAlaAsnLeuSerMetLysAlaTyrIleuGlyArgGlie 201
Db 617 TGGGTCATTCATCATCTGTTGGAATCTTGCAACTGACCCGGAATGAAGCTAGAAAC 676
QY 202 ArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrIleuArgThrGlnCys 221
Db 677 AGAGACAGTGTATGCAATCATGATGATGATGATGATGATGATGATGATGATGATG 736
QY 222 GlyGlnLeuTyrAlaTyrAspTyrPheSerIleProLeuValTyrThrGlnValValThr 241
Db 737 AGCCTTATTCGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 787
QY 242 ValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnProAsn 261
Db 787 ----- 787
QY 262 LysAspTyrProGlyHisGlnMetAspLeuValProValPheThrIleLeuGlnPhe 281
Db 787 ----- 787
QY 282 LeuPheTyrMetGlyTTPLeuLysValAlaGlnGlnLeuIleAsnProPheGlyLysP 301
Db 788 -----GTAGCAGACGAGCTTATTCACCCCTTTTGGACAGAT 823
QY 302 AspAspAspPheGluThrAsnTyrPheIleLeuAspArgAsnLeuGlnValSerLeuLeuSer 321
Db 824 GATGATGATTTTGAACCTAAGCTGGTCATGTGACAAATTTGACAGCTCTCTTTAGCT 883
QY 322 ValAspGlyMetHisGlnAsnLeuProPheMetGlnArgAspMetTyrTTPAsnGlnAla 341
Db 884 GTGAGCGAAATGCAATGATGCTTACCCAAAGATGAAGAGCATTTACTGGGAGATTC 943
QY 342 AlaProGlnProPheTyrThrAlaAlaSerAlaArgSerArgArgHisSerPheMetGly 361
Db 944 GCTGCTGCCCAACCATACATGTGACAGCTGCTGACTACTGATACCTCATTTCTGCGG 1003
QY 362 Ser 362
Db 1004 TCA 1006

```

RESULT 2

```

US-09-768-826-16
Sequence 16, Application US/09768826
Patent No. US20020012966A1
GENERAL INFORMATION:
APPLICANT: Shi et al.
TITLE OF INVENTION: 18 human secreted proteins
FILE REFERENCE: PF512P1
CURRENT APPLICATION NUMBER: US/09/768,826
CURRENT FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: PCT/US00/22350
PRIOR FILING DATE: 2000-08-15
PRIOR APPLICATION NUMBER: 60/148,759
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 16
LENGTH: 1350
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1135)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1209)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1229)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1266)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1285)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1287)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1290)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1295)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1305)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1324)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c

```

NAME/KEY: SITE
 LOCATION: (1345)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: SITE
 LOCATION: (1348)
 OTHER INFORMATION: n equals a,t,g, or c
 US-09-768-826-16

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5	84e-65	1350	641.50	147	42	86	76
Percent Similarity:	53.85%						
Best Local Similarity:	41.88%						
Query Match:	21.79%						
DB:	10						

US-09-622-964-29 (1-551) x US-09-768-826-16 (1-1350)

QY 179 AsnThrPheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrLeuGly 198
 DB 59 AACAGTACTGGTGGTCCCTGCTGCTGTTACACCACTGGCGCCGAGCCGAGGAGC 118
 QY 199 GlyArgLeuArgAspThrValLeuLeuGlnSerLeuMetAsnGluValCysThrLeuArg 218
 DB 119 GGGCGAATACGTCGACGATATCGCTCTGCTACTTTTGAAGAGCTGACAGTACCGA 178
 QY 219 ThrGlnCysGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGln 238
 DB 179 GCCAAGTGCACCAATGCTATTCACATGACTGATCAGATCAACCCCTGCTACACCCAA 238
 QY 239 ValValThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeu 258
 DB 239 GTGTGACATACATAGCGCTACTCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 298
 QY 259 AsnProAsnLysAsp-----TyrProGlyHisGlu----- 268
 DB 299 GAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 358
 QY 269 -----MetAsnLeuValValProValPheThrIleLeuGlnPheLeuPhe 283
 DB 359 CCAGCCCTGGAGAGCCCGGACATGTACGTGCTCTGACACCATCTGCTGCTGCTGCTGCT 418
 QY 284 TyrMetGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGluAspAsp 303
 DB 419 TATGCTGGCTGCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418
 QY 304 AspPheGluThrAsnTrpIleLeuAspArgAsnLeuGlnValSerLeuLeuSerValAsp 478
 DB 479 GACTTTGAGACAAATACATCATATGACGCAACCTTGCAGGTGCTCCCTGCTATCCGTGAC 538
 QY 324 GlyMetHisGlnAsnLeuProPheMetGluArgAspMetTyrTrpAsnGluAlaAlaPro 343
 DB 539 GAATGTACAGAACTTCCCTCCGCTGAGAGAGGACAGTACTGGATGAGGACCAAGCCG 598
 QY 344 GlnProProTyrThr-----AlaAlaSerAlaArgSerArgTrpHisSerPheMetGlySer 362
 DB 599 CAGCCACCTACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 658
 QY 363 ThrPheAsnIleSerLeuLysGluAspLeuGlnLeuTyrPheLysGluGluAlaAsp 382
 DB 659 ACCCTCAACTCGGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 718
 QY 383 ThrAspLysLysGlnSerGlyTyrSerSerThr-----IleGlyCysPheLeuGlnLeu 400
 DB 719 GGATCTGTGCTGGCGCCGCGCGCGGAGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 778
 QY 401 GlnProLysAsnThrHisLeuProLeuLysAspLeuLys-ThrLysLeuLeuCysSerLys 420
 DB 779 GGGGGGCGCTCC-----CCGGCATACAGCTTCGGAATTTGGCGCGGTGCGAGC 829
 QY 420 AsnProLeuLeuGlnGlyGlnCys----- 428
 DB 830 ACCCCCGCCCGCCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889

QY 429 -LysAspAlaAsnGlnLysAsnGlnLysAsp-----ValTyrLysPhe 442
 DB 890 GGAGCCCGCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 438
 QY 439 -----ValTyrLysPhe 442
 DB 950 TTACAGCCCTGTTTCATGTGTGAGAACGCTCATCTGTGAAGTGTGTGAAGTGTGAAGT 1009
 QY 442 eLysGlyLeu-----AspPheLeuLys----- 449
 DB 1010 GGATGGATTTCAAGACTTTAGAGAAAGAACCCACTGTTTATGACAGATCTTGCTCA 1069
 QY 450 -Cys-----ValProArgPheLysArg 456
 DB 1070 CTGCACCTCCACTCCCGAGGTTCAAGAGA 1100

RESULT 3

US-09-746-783-3

Sequence 3, Application US/09746783
 Publication No. US20030044935A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John M.

Lavallee, Edward R.

Racie, Lisa A.

Tracy, Maurice

Spaulding, Yikl

Agostino, Michael J.

Hoves, Steven H.

Feichtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESS: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/746,783

FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.

REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1263 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-746-783-3

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3	46e-61	1263	609.00	140	37	74	50
Percent Similarity:	58.80%						
Best Local Similarity:	46.51%						
Query Match:	20.69%						
DB:	12						

US-09-622-964-29 (1-551) x US-09-746-783-3 (1-1263)

```
OY 292 GUGLHLeuIleAsnProheglYluAspAspAspPheglUthAsnTrpIlelle 311
Db 1 GAGCAGCTCATCAACCCCTTGGAGAGATGATGATTTGAGACCAACAGCTGCTTC 60
OY 312 AsPArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGlnAsnLeuPro 331
Db 61 GACAGGAATTTGCAGGTGCTGCTGTGGCTGTGATGATGATGATGATGATGATGAT 120
OY 332 MetGluAspAspMetTrpTrpAsnGlnAlaAlaProGlnProProTrpTrpAlaIle 351
Db 121 ATGGAGCCGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
OY 352 AlaArgSerArgHisSerPheMetGlySerTrpPheAsnIleSerLeuLysGln 371
Db 181 GCCCAGTTCCTCGAGGCTCTCTTATGGCTCCACCTTCAACATCAGCTGACAAAG 240
OY 372 AsPLeuGluLeuTrpSerLysGlnGlnAlaAspTrpAspLysGlnSerGlyTrp 391
Db 241 GAGATGAGATTCACAGCCCAATCAGAG-----GACGAGAGATGCTCAGCCT 288
OY 392 SerTrpIleGlyCysPheLeuGlnLeuGlnProLysAsnTrpHisLeuProLeuLys 411
Db 289 GGCATCATGGCCCTTCTCTAGGCTCAGTCCATCATCATCATCTCTCCAGGGCAAC 348
OY 412 LeuLysTrpLysLeuLysSerLysAsnProLeuLeu-----GluGlnGlnCysLys 429
Db 349 TCAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCCACAGAGGCTGCCCAA 408
OY 430 Asp-----AlaAsnGlnLysAsnGlnLysAsp-----ValTrpLysPhe 442
Db 409 AACCAACAAGCAGCAAAACAAGAGTTAGGGCCAGCAACAACAAGAGCTGAGAGCT 468
OY 443 LysGlnLysAspPheLeuLysCysValProArgPheLysArgLysSerHisCysGly 462
Db 469 AAGCTGTGGAGCGCTTCAAGTCTGCGCCCATCTATCAGAGCCAGCTACTACAGTGC 528
OY 463 ProGlnAlaProSerSer-----HisProTrpGlnInsSerAlaProSer 477
Db 529 CCACAGAGCCCTCAGCCCTCCAGCTCCAGTGTCTCCCTGAGAACATCAGAGCGCTCA 588
OY 478 Ser-----SerAspTrpGly----- 482
Db 589 AAGCTTCAAGTCTCAGAGCATAGACACCAAGAACAAAGCTTAAAGCTGTGAGTCT 648
OY 483 -----AspGlyProSerTrpAspTrpGln 490
Db 649 GGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGCGGAGGCTGTGATGAGACCA 708
OY 491 GluIleCysHisMetLysLysTrpValGluPheAsnLeu-----AsnIleProGln 509
Db 709 GAAGTATCTCAAGTGAAGGAAACTGTGAGTTTAACTTACGAGATATGCAAGATC 768
OY 510 ProTrpGlnHisLeuGlnIleArgLysAspGlnMetSerTrpAsnIleGlnAlaLeu 529
Db 769 CCGGAAATCATCAAAAGAA-----CCTTGAACAATGACCAACACATACACACTCA 825
OY 530 MetLysGlnHisAlaGlnSerTrp-----ProTrpArgAspGlnAlaGlyTrpLys 546
Db 826 CTCAAAGATCATGATGATCTTATGCGCTTGGAAAACAGGATGAAGACATCTCTAA 885
OY 547 Pro 547
Db 886 CCT 888
```

RESULT 4
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James

```
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steilmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE OF INVENTION: THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR APPLICATION NUMBER: 2002-07-18
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11346
LENGTH: 853
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
LOCATION: 43, 50, 51, 54, 57, 853
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346
```

```
Alignment Scores:
Pred. No.: 2,51e-24 Length: 853
Score: 293.00 Matches: 54
Percent Similarity: 82.05% Conservative: 10
Best Local Similarity: 69.23% Mismatches: 14
Query Match: 9.95% Indels: 0
DB: 15 Gaps: 0
```

US-09-622-964-29 (1-551) x US-10-198-846-11346 (1-853)

```
OY 83 GlyPheYValTrpLeuValSerArgTrpSerGlnTrpGluAsnLeuProTrp 102
Db 273 GGGTTTATGTACTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 332
OY 103 ProAspArgLeuMetIleGlnValSerSerPheValGlnLysAspGlnGlnArg 122
Db 333 CAGACAGGCTAAAGTCTCTATCTCTACATCTCTACATCTCTACATCTCTCTCTCTCT 392
OY 123 LeuLeuArgArgTrpLeuIleArgTrpAlaIleLeuGlnGlnIleLeuValSer 142
Db 393 CTGCTTGAAGAGAGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 452
OY 143 IleSerTrpSerValTrpLysArgPheProTrpLeuHisIleValLeuAla 160
Db 453 GTGAGCACTGCTGTGTACAAAGATTTCCACAAATGAGACCACTGTGTGAACA 506
```

```
RESULT 5
US-10-027-632-134529/c
; Sequence 134529, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
```


PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 134529
 LENGTH: 748
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-134529

Alignment Scores:

Score:	2.01e-21	Length:	748
Percent Similarity:	267.50	Matches:	72
Best Local Similarity:	35.00%	Conservative:	5
Query Match:	32.73%	Mismatches:	13
DB:	9.09%	Indels:	131
	15	Gaps:	2

US-09-622-964-29 (1-551) x US-10-027-632-134529 (1-748)

```

QY 152 ProthLeuHisHisLeuValLeuAlaGlyPheMetThrHisGlyGluHisLysGlnLeu 171
DB 712 CCCACCGCTTTTCACTCCACTCTGCAGGCTTTATGACTCCGCGAACAACACAGCATG- 654
QY 172 GlnLysLeuGlyLeuProHisAsnThrPheTrpValProTyrValTrpPheAlaAsnLeu 191
DB 653 GAGAACTGAGCGCTACACACACATGTTCTGGGGCCCTGGGTGTGTTGCCAACCCTG 594
QY 192 SerMetLysAlaTyrLeuGlyGlyValArgLysAspThrValLeuLeuGlnSerLeuMet 211
DB 593 TCATGAGAGGGGTGGCTTGAGGTGAGTCCGAGACCTATCGTCTCCAGAGGCTCTG 534
QY 212 Asp----- 212
DB 533 AAGGTAGCCCACTGTACAGACAGGGGTGCCGAGATGGGAGGTTGGTCCACAGG 474
QY 212 ----- 212
DB 473 AAACAAGTTTCTACAAAGAGAAGCCTTGGGCCCTGAGGGTCTTCCGAGACCGGAGG 414
QY 212 ----- 212
DB 413 TGGGGTTCAGATCTTTTCCACAGCAATCCACAGCCCGAGGTGCTCTCTCAGAGG 354
QY 213 -----GluValCys----- 215
DB 353 CCCCTCCCTCTCTCCAAAGTCTGTGAGTCTGTCTGTTCCCTTTGATAGATGAGAACTG 294
QY 215 ----- 215
DB 293 AGACACAAGAGAGTTTGTAGACTTCCATGCGCCACACAGCCAGAGATGACATAGGTA 234
QY 215 ----- 215
DB 233 CCAGGCCCTGGTACTGTGAKAAGAGTGGGGGCGAGCCAGGTTGGGGGCGAGGTGTGTT 174
QY 216 -----ThrLeuArgThrGlnCysG1 222
DB 173 CAGAAACCCATCCCTCTTCTGCCCCCCCAGAGATGAAACACCTTGGTACTGCTGCG 114
QY 222 yGlnLeuTyrAlaTyrAspTrpLysSerIleProLeuValTyrThrGlnValTyr 241
DB 113 ACACCTGTATGCCCTACAGCATGATTAATCCCACTGGTGTATACACAGGTGAGACT 56

```

RESULT 6
 US-10-198-846-2561/c
 Sequence 2561, Application US/10198846
 Publication No. US20030099974A1
 GENERAL INFORMATION:
 APPLICANT: Lillie, James
 APPLICANT: Xu, Yongyao
 APPLICANT: Wang, Youzhen
 APPLICANT: Steinmann, Kathleen
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 FILE REFERENCE: THERAPY OF BREAST CANCER
 CURRENT APPLICATION NUMBER: US/10/198,846
 PRIOR FILING DATE: 2002-07-18
 NUMBER OF SEQ ID NOS: 14084
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2561
 LENGTH: 462
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 390, 462
 OTHER INFORMATION: n = A,T,C or G
 US-10-198-846-2561

Alignment Scores:

Score:	6.09e-20	Length:	462
Percent Similarity:	252.00	Matches:	46
Best Local Similarity:	82.09%	Conservative:	9
Query Match:	68.68%	Mismatches:	12
DB:	8.56%	Indels:	0
	15	Gaps:	0

US-09-622-964-29 (1-551) x US-10-198-846-2561 (1-462)

```

QY 82 LeuGlyPheTyrValThrLeuValValSerArgTrpSerGlnTyrGlnAsnLeuPro 101
DB 256 ATAGGCTTTATGTTACTCTGCTAGTGAACGATGGTGAACCACTTTGTGAATTTGCC 197
QY 102 TrpProAspArgLeuMetIleGlnValSerSerPheValGluGlyLysAspGluGly 121
DB 196 TGGCCAGACAGCTATATGTTCTCTATCTGCACTGTTCACCGAAGCAGACACAGG 137
QY 122 ArgLeuLeuArgArgThrLeuIleArgTyrAlaIleLeuGlyGlnValIleLeuArg 141
DB 136 CGCTGCTTAAGAGACCGCTGATCGGTACGTCATATCCACCTCCCTGCTCATCTTTCG 77
QY 142 SerIleSerThrSerValTyr 148
DB 76 TCGGTGAGCACTGCTGTGTAC 56

```

RESULT 7

US-10-198-846-7843

Sequence 7843, Application US/10198846
 Publication No. US20030099974A1
 GENERAL INFORMATION:
 APPLICANT: Lillie, James
 APPLICANT: Xu, Yongyao
 APPLICANT: Wang, Youzhen
 APPLICANT: Steinmann, Kathleen
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
 FILE REFERENCE: THERAPY OF BREAST CANCER
 CURRENT APPLICATION NUMBER: US/10/198,846
 PRIOR FILING DATE: 2002-07-18
 NUMBER OF SEQ ID NOS: 14084
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 7843
 LENGTH: 615
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-198-846-7843

Alignment Scores:

Score:	9.7e-20	Length:	615
Pred. No.:	252.00	Matches:	46

Percent Similarity: 82.09%
Best Local Similarity: 68.66%
Query Match: 8.56%
DB: 15
Conservative: 9
Matches: 12
Indels: 0
Gaps: 0

US-09-622-964-29 (1-551) x US-10-198-846-7843 (1-615)

OY 82 LeuGlyPheTyValThrLeuValIserArgTrpSerGlnTyrGlnAsnLeuPro 101
DB ATAGGTTTATCTACTCTGAGTGAACCGATGGTGGACCGAGTTGTGAATTTGCC 411
OY 102 TTPProAspArgLeuMetIleGlnValSerSerPheValGluGlyLysAspGluGly 121
DB 412 TGGCCAGACAGGCTATGTTCTCTACAGTCTACAGTGTCCAGGAGACAGACAGCGG 471
OY 122 ArgLeuLeuArgThrLeuLeuLeuGlyValIleLeuGlyGlnValLeuLeuArg 141
DB 472 CGCGCTGTAAGAAAGCGCTGATGCGCTACGTAATCTCACCTCTCATCTTTGCG 531
OY 142 SerIleSerThrSerValTyr 148
DB 532 TCGGTGAGCAGCTGTGTGATC 552

RESULT 8

US-10-027-632-134530/c
Sequence 134530, Application US/10027632.
GENERAL INFORMATION:

APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530

Alignment Scores:

Prod. No.: 2,59e-19 Length: 578
Score: 248.00 Matches: 46
Percent Similarity: 81.97% Conservative: 4
Best Local Similarity: 75.41% Mismatches: 11
Query Match: 8.42% Indels: 0
DB: 15 Gaps: 0

US-09-622-964-29 (1-551) x US-10-027-632-134530 (1-578)

OY 152 ProThrLeuGlnIleValLeuValIleGlyPheMetTrpHisGlyGlnHisLysGlnLeu 171
DB 207 CCACGCGCTTCTTCACTCCACCTGTCAGGCTTTATGACTCCGCGAAGACACACAGCTTG 148
OY 172 GlnLysLeuGlyLeuProHisAsnThrPheTrpValProTrpValTTPPheAlaAsnLeu 191
DB 147 GAGAACTGAGGCTACACACAAACATGTTTGGGGGCGGCTGGGTGGTTGGCCAACTG 88

OY 192 SerMetLysAlaTyrLeuGlyArgIleArgAspThrValLeuLeuGlnSerLeuMet 211
DB 87 TCATGAAAGGCGTGGCTTGAGTGCATCCGGACCCATCTCTGCTCCAGACCTGCTG 28

OY 212 Asn 212
DB 27 AAC 25

RESULT 9

US-09-764-872-195
Sequence 195, Application US/09764872
Publication No. US20030050231A1

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P125
CURRENT APPLICATION NUMBER: US/09/764,872
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 957
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 195
LENGTH: 507
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (490)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-872-195

Alignment Scores:

Prod. No.: 1,59e-17 Length: 507
Score: 232.00 Matches: 46
Percent Similarity: 73.49% Conservative: 15
Best Local Similarity: 55.42% Mismatches: 22
Query Match: 7.88% Indels: 0
DB: 12 Gaps: 0

US-09-622-964-29 (1-551) x US-09-764-872-195 (1-507)

OY 1 MethrIleuTrpTyrAsnLysValAlaAsnAlaGlnGlySerPheSerLeu 20
DB 37 ATGAGGTTTCAATACATCTCAAAAGTGGCGGAGCGCTTGGAGTTCTCTGGCGTG 96
OY 21 LeuGlyCysTrpArgLysSerIleTyrLysLeuLeuTyrGlyGlnPheValPheIle 40
DB 97 CTCTCCGCTGAGGAGGAGCATCTACAGCTCTCTACAAAGAAATTCCTCTTTGGG 156
OY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnLeu 60
DB 157 GCCTTGTACCTGCTTACATACCTACCTACCGGCTCTGTGACCCAGGAGCAGAGTAC 216
OY 61 LeuPheGlnLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuLeuProIleSerPhe 80
DB 217 GTGTATCTGACGTGGCGGCTGACTGCAACCGCTCAGCAGACTCATTCCTTGTCTTT 276
OY 81 ValLeuGly 83
DB 277 GTATTGGGT 285

RESULT 10

US-10-198-846-1862
Sequence 1862, Application US/10198846
Publication No. US2003009974A1

GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
TITLE OF INVENTION: STEINMANN, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER

Alignment Scores:		Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
			0.00899	120.50	33.39%	20.22%	5218	112	73	185	186	26
					4.09%							
US-09-622-964-29 (1-551) x US-10-225-567A-345 (1-5218)												
QY	69	CysAspSerTyrIleGlnLeuIleProIleSerPheValLeuGlyPheTyrValThrLeu	88	2854	TGCACACACACACACATGTCATTTTGGACCTTTTCTTCCTGGCTTCATTC	-----TGTTGG	290					
DB		111	111	111	111	111	111	111	111	111	111	111
QY	89	ValValSerArgTyrTrpSerGlnTyrGlnValLeuProTyrProAspArgLeuMetIle	108	2908	CTTTTGGACTGAGGCGCTGCATCATAT	-----	293					
DB		111	111	111	111	111	111	111	111	111	111	111
QY	109	GlnValSerSerPheValGlnGlyValAspGlnGlyArgLeuLeuArgArgThrLeu	128	2935	-----ATGGCTGTACTGGAATAATTAGACA	-----CGCTTATTAAGAAA	2976					
DB		111	111	111	111	111	111	111	111	111	111	111
QY	129	IleArgTyrAlaIleLeuGlyGlnValLeuIleLeuArgSerIleSerThrSerVal	147	2977	-----CGCTTTTGGCTCTTGATGGGTTTACACACATTAGTAGAGTGGCCACATCGTAGGC	3033						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	148	-----TyrIleArgPheProThrLeuHisHis	162	3034	TTCACACAGAACAAAGATATGGCACTGATACACTAGCTGCTCTCTTGAAGAGAGA	3093						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	163	MetThrHisGly	166	3094	CTACTCTATGCTTTTGTGGACCTGCAGCGCGTGTCTGTCTGTCACATGTGATTGCC	3153						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	167	-----	168	3154	ATTTTGGATTATTAATAACTGTTTCCACAGATGCAATCTAGATATAAAGCTCAACAC	3213						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	169	LysGlnLeuGlnLysLeuGlyLeuProHisAsnThrPheTyrValProTyrValThrPhe	188	3214	AGAGCGGCTGAG-----ATGAGTGAGCGCTCAT	-----	3240					
DB		111	111	111	111	111	111	111	111	111	111	111
QY	189	AlaAsnLeuSerMetLysAlaTyrIleGlyArgIleArgAspThrValLeu	207	3241	AGCGGTGTGAGCGTCAATATGTCACAGTGTGGAGTAGTTTCAACAAAGCTTTCAGAGCC	3300						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	208	GlnSerLeuMetAsnGlnValCysThrLeuArgThrGlnCysGly	225	3301	ACCCACGCCACAGTAAGCCATGGCTGCTCTTGGAGTCCCTGTGGTTCCTCCCTTTG	3360						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	226	AlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValValThrValAlaValTyr	245	3361	GCATTGACGAGATGTGTGGCGTTCTGGCCATGACAGATAAACGCTCCATAT	-----TGTTTC	3418					
DB		111	111	111	111	111	111	111	111	111	111	111
QY	246	SerPheLeuAlaCysLeuIle	253	3419	AAATACTTTTGGCTGTGATTGATTCAATTGCAAGGCTTGTATTAGTCATGTCACATGCA	3478						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	254	-----GlyArgGlnPheLeuAsnProAsnLysAspTyrProGlnHisGlnMetAspLeu	271	3479	TTCCTGGAGAGAGGCTTC	-----AGATGACATTTA	3508					
DB		111	111	111	111	111	111	111	111	111	111	111
QY	272	-----ValValProValPheThrIleLeuGlnPheLeuPheTyr	284	3509	GATGCCGATTGAGAAACTGTACAGATATCCATCAATGACATTTCTTGAGTTGCTTCCAA	3568						
DB		111	111	111	111	111	111	111	111	111	111	111
QY	285	MetGlyTyrPheLysValAlaGlnIleLeuAsnProPhe										


```

1  CONT: 7020LA, SERGEY
2  TITLE OF INVENTION: HUMAN OLFACATORY RECEPTORS AND GENES ENCODING SAME
3  FILE REFERENCE: P 0278005
4  CURRENT APPLICATION NUMBER: 2001-03-13
5  PRIOR APPLICATION NUMBER: 2001-03-13
6  PRIOR FILING DATE: 2000-03-13
7  PRIOR APPLICATION NUMBER: 60/198,914
8  PRIOR FILING DATE: 2000-03-13
9  PRIOR APPLICATION NUMBER: 60/192,033
10 PRIOR FILING DATE: 2000-03-24
11 PRIOR APPLICATION NUMBER: 60/198,474
12 PRIOR FILING DATE: 2000-04-14
13 PRIOR APPLICATION NUMBER: 60/199,335
14 PRIOR FILING DATE: 2000-04-24
15 PRIOR APPLICATION NUMBER: 60/207,702
16 PRIOR FILING DATE: 2000-05-26
17 PRIOR APPLICATION NUMBER: 60/213,849
18 PRIOR FILING DATE: 2000-06-23
19 PRIOR APPLICATION NUMBER: 60/226,534
20 PRIOR FILING DATE: 2000-08-16
21 PRIOR APPLICATION NUMBER: 60/230,732
22 PRIOR FILING DATE: 2000-09-07
23 PRIOR APPLICATION NUMBER: 60/266,862
24 NUMBER OF SEQ ID NOS: 529
25 SOFTWARE: PatentIn Ver. 2.1
26 SEQ ID NO 22
27 LENGTH: 930
28 TYPE: DNA
29 ORGANISM: Homo sapiens
30 US-09-804-291-22

```

Alignment Scores:	
Pred. No.:	0.0471
Score:	104.00
Percent Similarity:	37.06%
Best Local Similarity:	24.48%
Query Match:	3.53%
DB:	12
	Gaps:
	12

[illegible]

```

Qy      128 -----Le 121
Db      528 TCTGTGTACGCTTCTCCCTCTTTCAGACTCTCTCCACCAACCCATGTCAGTACT 11
Qy      128 CILGARGTVALALEULEUGLYINLVALLEULLEUARGSERILESERTHSERVALTY 146
Db      588 GGTATTTTCTCTGTCTGTGGATATTCACCATGCTATGACACATACCATACATGCTGATC
Qy      148 TLVARGPHEPRTHTREUNISHISEUVALLEUALGLYPHEMETHRHSGLYUHI 168
Db      648 -TTACGCTGTGATACCTCCACACATCTCTGATTC-----TTCTGCAGAGGGCAGT 700
Qy      168 SLVSGINLEUGINLVLEUGLYLEUPROHISASN-----ThrethrpValPr 184
Db      701 CCAACGCTTTTGGACATGAGGGCTCCACATATTTGCGTGTGCTCTGTTTTTGGGCTC-- 758
Qy      184 OTTPVALTRPHEALASNLSEURMELLYSALATYRLEUGLYARGILEARGSPH 204
Db      759 -----AGGGACAT 766
Qy      204 rValLeuLeuGlnSerLeuMetAsn-----GluValCys---ThrLe 217
Db      767 TCACCTACTTAAACAACATCTTTCTGCTGCTCTATGACCACTGGACAGATTTGCTAGTCT 826
Qy      217 uArgThrGlnCysGly 222
Db      827 TTTACCAACCAATGTGT 842

```

```

RESULT 14
US-10-027-632-245664/C
/ Sequence 245664: Application US/10027632
/ GENERAL INFORMATION:
/ APPLICANT: Rang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 106827.129 Polymorphisms in the Human Genome
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/199,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 245664
/ LENGTH: 609
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-245664

Alignment Scores:
Pred. No.:
Score: 0.0406 Length: 609
Percent Similarity: 102.00 Matches: 18
Best Local Similarity: 77.14% Conservative: 9
Query Match: 3.46% Mismatches: 8
DB: 15 Indels: 0
Caps: 0

US-09-622-964-29 (1-551) x US-10-027-632-245664 (1-609)
OY 51 ArgmetValIeuSerSerAspGlnGlnIleuLeuPhcGluLysIreuAlaIeuTyrCysAsp 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

Db 370 AGATTGTTACTTACAGAGGACCAAAACGTTACTTGAATAATATATCAATTACTGTGAC 311
 QY 71 SerYrILeGLInLeuILeProILeSerPheValLeuGLYpHeTYr 85
 Db 310 AGATATGCTGAAACAATTCAGTAACCTTGTGCTGTGTAAGTAT 266
 RESULT 15
 US-09-750-373-12
 ; Sequence 12, Application US/09750373
 ; Patent No. US20020062013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lind, Peter
 ; APPLICANT: Wood, Linda S.
 ; APPLICANT: Hiebsch, Ronald
 ; APPLICANT: Ruff, Valerie
 ; APPLICANT: Lindberg, Eleni
 ; APPLICANT: Parodi, Luis A.
 ; APPLICANT: Vogel, Gabriel
 ; TITLE OF INVENTION: No. US20020062013A1 G Protein Coupled Receptors
 ; FILE REFERENCE: PHM-0300
 ; CURRENT APPLICATION NUMBER: US/09/750,373
 ; CURRENT FILING DATE: 2000-12-28
 ; PRIOR APPLICATION NUMBER: 60/184,305
 ; PRIOR FILING DATE: 2000-02-23
 ; PRIOR APPLICATION NUMBER: 60/188,880
 ; PRIOR FILING DATE: 2000-03-13
 ; PRIOR APPLICATION NUMBER: 60/219,492
 ; PRIOR FILING DATE: 2000-07-20
 ; PRIOR APPLICATION NUMBER: 60/173,339
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/224,321
 ; PRIOR FILING DATE: 2000-08-11
 ; PRIOR APPLICATION NUMBER: 60/200,534
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 60/239,062
 ; PRIOR FILING DATE: 2000-10-09
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 12
 ; LENGTH: 1668
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-750-373-12
 Alignment Scores:
 Pred. No.: 0.21 Length: 1668
 Score: 102.00 Matches: 108
 Percent Similarity: 35.09% Conserved: 85
 Best Local Similarity: 19.64% Mismatches: 187
 Query Match: 3.46% Indels: 170
 DB: 10 Gaps: 29
 US-09-622-964-29 (1-551) x US-09-750-373-12 (1-1668)
 QY 2 ThrILeThrTYrThrAsnLYsValAlaAsnAlaArgLeuGLYSerPhe----- 17
 Db 213 ACAAGTGCCTTTACTAGAGTGGAGGAAGAAAGAAAGGAGTTCCTTACTACTCCTTT 272
 QY 18 -----SerSerLeuLeuLeuCYsTrpArgGLYSerILeTYrLYsLeuLYrGLYGLu 35
 Db 273 AAGACGTGACGAATGTGATACCTGTGGGCTCTTGTCTTTTACCATGTGTGGAACCTCC 332
 QY 36 PheLeuValPheILePheLeuTYrTYrSerILeArgGLYLeuTYrArgMetValLeuSer 55
 Db 333 GTTGCTCTTTT-----TCCACATGGAGGAGAAAGAAAG 368
 QY 56 SerAspGLInLeuLeuPheGLYLYsLeuAlaLeuTYrCYsAspSerTYrILeGLInLeu 75
 Db 369 TCAAGATGACCTTCTTTGTGACTGACGTGGCCATC--ACAGATTCTTTCACAGAGACTG 425
 QY 76 ILeProILeSerPheValLeuGLYpHeTYrValThrLeuValValSerArgTrpTrpSer 95
 Db 426 GTCAACATC-----TTGACAGATATATATTGGGATTCACCTGGA 464

QY 96 GLNLYrGLInAsnLeuProTrpProAspArg----- 105
 Db 465 GACTTCACAGGCA-----CCTGACCTGGTTTCCGAGTGTGCCGTATTTCAGAGTT 515
 QY 106 LeuMetILeGLInValSerPheValGLInGLYLYsAspGLInGLYArgLeuLeuArg 125
 Db 516 GTGCTGCTCTAGCCCTTCACTACCTAGTC-----CTGTGTCC 551
 QY 126 ArgThrLeuILeArgTYr--AlaILeLeu----- 134
 Db 552 CTCAGCATAGACAGATACCATGCTGATCCCATCCCATGAGTTCCTTCAAGAGCAAG 611
 QY 135 --GLYGLInValILeLeuArgSerILeSerThrSerValTYrLYsArgPheProThr 153
 Db 612 CAAGCCAGGCTCTCATTTGTGATGCTGAGCCCTGAGCCCTGTTCTGTTCTCCATCCACC 671
 QY 154 LeuHISLeuValLeuAlaGLY-----PheMetThrHISGLYHISLYSGLInLeu 171
 Db 672 -----CTGATCATATTTGGGAAGAGCACTGCCAACGGTGAA-----GTG 713
 QY 172 GLNLYsLeuGLYLeu--ProHISAsnThrPheTrpValProTrp----- 185
 Db 714 CAGTGTGGGCGCTGTGGCCCTGGCGACTCCTGACCCCATACATGACCATCGTGGCC 773
 QY 186 -----ValTrpPheAlaAsnLeuSerMetLYsAlaTYrLeuGLYArgILeArgAsp 203
 Db 774 TTCCTGTGTACTTCATCCCTCTGACATATACGATCATGTATGCGATTCGATCCGA 833
 QY 204 ThrValLeuLeuGLInSerLeuMetAsnGLYValCYsThrLeuArgThrGLYInCYsGLYGLN 223
 Db 834 ACTATTTGGATTTAAAGCAAAACCTACGAACAGTATTTCCAACTGCTCAGATGGGAA 893
 QY 224 Leu-----TYrAlaTYrAspTrpILeSerILeProLeuValTYrThrGLNValVal 240
 Db 894 CTGTGACGACGATATACCGAGACATCATCTCAAGCAAAATCAAGGCTATCAAGTAT 953
 QY 241 ThrValAlaValTYrSerPhePheLeuAlaCYsLeuILeGLYArgGLInPheLeuAsnPro 260
 Db 954 AGCATATATCATCTTCTGCTTCATCTGCTG-----TGAGTCCA 995
 QY 261 AsnLYsAspTYrProGLYHISGLInMetAspLeuValProValPheThrILeLeuGLN 280
 Db 996 -----TAC 998
 QY 281 PheLeuPheTYrMetGLYTrpLeuLYsValAlaGLInLeuILeAsnProPheGLY--- 299
 Db 999 TTCCTGTGTT-----GACATTTTGGACAATTTCAACCTC 1031
 QY 300 -----GLYAspAspAspAspPheGLInThrAsnTrpILeILeAspArgAsnLeuGLNVal 317
 Db 1032 CTTCACAGACACCCAGGAGCGTTTATGCTCTGTGATCAAT----- 1073
 QY 318 SerLeuLeuSerValAspGLYMetHISGLInAsnLeuProPheMetGLYArgAspMetTYr 337
 Db 1074 -----CAGAACCTGCCAGCATTG----- 1091
 QY 338 TrpAsnGLInAlaIleProGLInProProTYrThrAlaAsnSerAla----- 352
 Db 1092 ---AATAGTGCATCAACACCCCCCATCATGCTGTCTTACACACCTCATCTCTTCCCC 1148
 QY 353 ---ArgSerArgArgHISSerPheMetGLYSerThrPheAsnILeSerLeuLYsLYsGLu 371
 Db 1149 TGCAGGGAGCAAAATATACAGAGATTCAGAAATGACGTTCCGGGAGAGAACTGAGAGGCAT 1308
 QY 372 AspLeuGLInLeuTrpSerLYsGLInGLN----- 380
 Db 1209 GAGATGAGATTCGTGTCAGGCCAGAAATCATATACAGACCTTAGGCGCATGCTGCTAG 1268
 QY 381 AlaAspThrAspLYsLYsGLYSerGLYTYrSerSerThrILeGLYCYsPheLeuGLYLeu 400
 Db 1269 GCTGACGACATCATGCTCTCCACAGGTCCTTGTACCTGCTTGGGACAGTGCATGGAACCC 1328

```

QY 401 GlnProLysAsnTyrHisLeuPro-----LeuLysAspLeuLysThrLysLeuLeu 417
Db 1329 GAGCAACTTCACCCAC---CCTGCTATTACTGGAGATGCACAAAGACAAA----- 1379
QY 418 CysSerLysAsnProLeuLeuGlnGlnCysLysAspAlaAsnGlnLysAsnGlnLys 437
Db 1380 TGTCTAATGACTGCATG-----CACTGCTTAAGTATGGCCAAACACGAA----- 1424
QY 438 AspValTrpLysPheLysGlyLeuAspPheLeuLysCysValProArgPheLysArgArg 457
Db 1425 -----CTCCCAAGTATTATCATGCACGCCAGAGAAAGGCCCTT 1463
QY 458 GlySerHisCysGlyProGlnAlaProSerSerHisProThrGlnGlnSerAlaProSer 477
Db 1464 CCTTCC-----CCACCATTCACACCCCTCCTCCCACTGGCCAGCACACTGAACCA 1514
QY 478 SerSerAspThrGlyAspGlyProSerThr 487
Db 1515 GTGAACACAGGCATTAGTGGTCCAGGGCTCC 1544

```

Search completed: July 27, 2003, 20:50:06
 Job time : 340.882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: July 27, 2003, 11:08:55 ; Search time 68.042 Seconds
(without alignments)
2483.449 Million cell updates/sec

Title: US-09-622-964-29
Sequence: 1 MTTTNNKVVANRIGSRSSL.....EHAESYPRDEAGTKPFLYE 551

Scoring table:
BLOSUM62
Xgapop 10.0 / Xgapext 0.5
Ygapop 10.0 / Ygapext 0.5
Delop 6.0 / Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=x1h
-DB=us09-622-964-29/US09622964/runat_22072003_101119_26036/app_query.fasta_1.2069
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09622964.GCN.1.1-79-@runat_22072003_101119_26036 -NCPU=6 -ICPU=3
-NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=10 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents, NA: *
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
4: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *
5: /cgn2_6/ptodata/1/lna/5A_COMB.seq: *
6: /cgn2_6/ptodata/1/lna/5B_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	3.8	2138	4	US-09-276-531-22
2	100	3.4	4803	4	US-09-197-636-1
3	100	3.4	4803	4	US-09-197-636-1
4	99.5	3.4	2544	2	US-08-463-412A-6
5	99.5	3.4	2544	4	US-09-021-715-6
6	99	3.4	3463	4	US-09-533-220A-1
7	97.5	3.3	3500	4	US-09-197-636-7
8	97	3.3	1230	3	US-09-188-930-25
9	97	3.3	1230	3	US-09-188-930-25
10	96.5	3.3	2544	4	US-09-188-930-201
11	95.5	3.2	2749	1	US-09-235-451-33
12	94.5	3.2	3802	6	US-07-717-331F-4
					Patent No. 5386025

13	94	3.2	13926	4	US-08-961-527-5	Sequence 5, Appl
14	93	3.2	3980	1	US-08-233-008A-1	Sequence 1, Appl
15	93	3.2	3980	1	US-08-233-008A-5	Sequence 1, Appl
16	93	3.2	6519	1	US-08-233-008A-7	Sequence 7, Appl
17	93	3.2	9334	4	US-09-172-952-19	Sequence 19, Appl
18	92.5	3.1	3387	4	US-08-468-557-1	Sequence 1, Appl
19	92.5	3.1	6305	4	US-09-221-017B-256	Sequence 256, App
20	92	3.1	2887	4	US-09-183-253-1	Sequence 1, Appl
21	91	3.1	5720	4	US-09-442-100-1	Sequence 1, Appl
22	91	3.1	8478	3	US-08-676-818-1	Sequence 1, Appl
23	91	3.1	8478	4	US-09-407-549-1	Sequence 1, Appl
24	90.5	3.1	2993	2	US-08-415-593-42	Sequence 42, Appl
25	90	3.1	2055	4	US-09-134-078-58	Sequence 58, Appl
26	90	3.1	3545	2	US-08-663-112-1	Sequence 1, Appl
27	89.5	3.0	2880	4	US-09-235-451-1	Sequence 1, Appl
28	89.5	3.0	7493	1	US-08-212-133A-7	Sequence 1, Appl
29	89.5	3.0	7493	2	US-08-474-503-5	Sequence 5, Appl
30	89.5	3.0	7493	2	US-08-670-707A-5	Sequence 5, Appl
31	89.5	3.0	7493	4	US-09-037-601-5	Sequence 5, Appl
32	89.5	3.0	7493	4	US-09-315-119-5	Sequence 5, Appl
33	89.5	3.0	7493	4	US-09-037-601-5	Sequence 5, Appl
34	89.5	3.0	2880	2	PCT-US94-13200-5	Sequence 5, Appl
35	89.5	3.0	2880	3	US-08-592-874-1	Sequence 1, Appl
36	89.5	3.0	2880	3	US-09-096-867-2	Sequence 2, Appl
37	89.5	3.0	2880	3	US-09-453-702B-60	Sequence 2, Appl
38	89	3.0	49795	4	US-08-874-138-1	Sequence 1, Appl
39	89	3.0	2800	2	US-08-874-138-5	Sequence 5, Appl
40	89	3.0	2800	4	US-08-879-941-1	Sequence 1, Appl
41	89	3.0	2800	4	US-08-879-941-3	Sequence 3, Appl
42	89	3.0	2800	4	US-09-747-116-1	Sequence 1, Appl
43	89	3.0	2800	4	US-09-747-116-1	Sequence 1, Appl
44	88.5	3.0	1851	4	US-08-965-902A-1	Sequence 1, Appl
45	88.5	3.0	1961	4	US-09-509-902A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-276-531-22
Sequence 22, Application US/09276531
Patent No. 6183968

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yee, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Bauglin, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276, 531
FILING DATE: Herewith

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079, 677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

379 GIUGLIAAspThrAspPylsGluSerClyTyrSerSerThrIleGlyCysPheLeu 398
||||| |||::: :: |||||::: ||| :::

REFERENCE/DOCID NUMBER: 6
TELECOMMUNICATION INFORMATION

TELEPHONE: 601-407-0700
 TELEFAX: 610-407-0701
 TELEX: 846169
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4803 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-197-636-1

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0.941	100.00	4803	103	61	193	160	23
Percent Similarity:	31.72%						
Best Local Similarity:	19.92%						
Query Match:	3.40%						

US-09-622-964-29 (1-551) x US-09-197-636-1 (1-4803)

```

OY 26 GlycerilLeuTyLysLeuLeuTyrelglnPheLeuValPheLeuTyrriser 45
Db 2292 GGAGAGATCTCTCTGTTAGAGAGG-----GTCACCTCTTT 2330
OY 46 lIeargLyLeuTyArgmetValLeuSerSeraspGlnGlnLeuPheglnLysLeu 65
Db 2331 TTCGAGGATTCACATATTTCTCTG-----CAGAGCGCGCCCTCATGAG 2375
OY 66 AlaLeuTyrcysaspSerTyrrileGlnLeuLeuProlleSerPheValLeuGlyPheTy 85
Db 2376 ACCCTGTTGTGGACAGCTACAGTACATG-----CTTCTCTTCTGCACTACCTGTC 2429
OY 86 ValrLeuValLeuSerArgTrpTrpSerLlnTyrlGlnasn----- 99
Db 2430 ATGCTGGCCACCGGGGCTGCTACTTCAGCCACCTCAGAGTATGTGCTTCATGTA 2489
OY 100 -----LeuProTrpProaspArgLeuMetIleGlnValSerSerPheValGlu 115
Db 2490 TTCCTCCCTGCGCTGGCTGACCATCATGCTCTAC-----TACACCCGC 2534
OY 116 GlyLysaspGlnGlnGlyArgLeuLeuArgTrpLeuLeuArgTyrralaIleLeuGly 135
Db 2535 GGTTCACAGACAGATGGGC-----ATCATGCGCCATGATA 2570
OY 136 GlnValLeuLeuLeuTyrriserThrSerValTyrralrGhrPheProThrLeuHis 155
Db 2571 GAGAGATGATCTCTGAGAGACCTGTC-----CGTTTCATGTTTGTCTAC 2615
OY 156 HisLeuValLeuLaglyPheMetThr-----HisGlyGlnHis 168
Db 2616 ATCGCTTCTTGTTCGGGTTTTCACAGCGGTGTCAGCTCATGTAAGACGGAGAGAT 2675
OY 169 LysGlnLeuGlnLysLeuGlyLeuProHisasnThrPheTrpValProTrpValrPhe 188
Db 2676 GACTCCCTGCGCTGAGTCCACGTCGACAG----- 2708
OY 189 AlaasnLeuSerMetLysAlaTyrrileGlnGlyArgIleArgspThrValLeuGln 208
Db 2709 -----TGCGGGGGCGCTGCTCAGAGCCCC-----GAT 2738
OY 209 SerLeuMetasnGlnValCysThrLeuArgThrGlnCysGlnLeuTyrrala----- 226
Db 2739 AGCTGCTCAAC-----AGCTGTACTCCACCTGCGGAGCTGTTCAAGTTCACC 2789
OY 227 -----TyraspTrpLieserLleProLeuVal 235
Db 2790 ATCGCATGGCGCATGAGATTACAGAGACATATGACTTCAAGCTGCTTCATCATC 2849
OY 236 TyrThrGlnValValThrValAlaValTyrrSerPhePheLeuValLacysLeuGlyArg 255
Db 2850 CTGCTGCTGGCCTATGTAATTCACCTCATCTCTGCTCAACATGCTCATCGCCCTC 2909

```

```

OY 256 GlnPheLeuasnProasnLysaspTrpProGlyHisGln----- 268
Db 2910 ATGGGTGACACTGCTCAACAAGATCCGACAGAGACAGACATCATCTGAGAGCTGCAGAGA 2969
OY 269 -----MetaspLeuValValProValPheThrIleLeuGlnPheLeuPheTy 284
Db 2970 GCCATCACCATTCTGGACAGGAGAGAGCTTCTTAAGTCATGAGAGAGCTTCCCGC 3029
OY 285 MetGlyTrpLeuLysValAlaGlnLeuLeuLeasnPropheGlyGlnLysaspasp 304
Db 3030 TCAGGC-----AACCTGCTGACAGGTGGGTACACACCTGATGGCAGAGACTACCGC 3083
OY 305 Phe-----GluThrAsnTrpLle----- 310
Db 3084 TGTGCTTCAGAGGTGAGACAGATGAACTGACACCTGGAAACACCAACGTCGATCATC 3143
OY 311 -----IleaspArgasnLeuGlnValSerLeuLeuSer 321
Db 3144 AACGAGACCGCGCAACTGTGAGGCGGTCAAGCGCACCTGACTTCTCCTGCGGTCA 3203
OY 322 -----ValaspGlyMetHisGlnasnLeuProPheMetLuaraspMetTyrrPasn 339
Db 3204 AGCAGACTTTCAGGACAGAC-----TGGAAG 3230
OY 340 GlnAlaAlaProGlnProProTyrrThrAlaAlaSerAlaArgSerArgArgHisSerPhe 359
Db 3231 AACTTGGCCCTGCTCCCTTTAAGAGAGGCAAGTCTCGAATAGGCACTGCTCAG 3290
OY 360 MetGlySerThr-----PheasnLieserLeuLysGlnaspLeuGln 375
Db 3291 CCGGAGAGAGTTATATCTGACAGACTTTCAGGCTCTCTGACACAGAGACCTGAGTGC 3350
OY 376 TrpSerLysGlnGlnAlaaspThrAspLysLysGlnSerGlyTyrrSerThrIleGly 395
Db 3351 TTTCAAGATCTGCGCTCTCCGGGAGAGTACAGACGTCACGACAGACACTGTC--- 3407
OY 396 CysPheLeuGlyLeuGlnProLysasnTyrrHisLeuProLeuLysaspLeuLys---Thr 414
Db 3408 -----AACACTGGGCTTAAAGAGACCCCGTGGCCACAGGGGGGCTGTCAGGAGACA 3458
OY 415 LysLeuLeuCysSerLysasnProLeuLeuGlnGlyGlnCysLysaspAlaasnGlnLys 434
Db 3459 CCAGTGCCTGT---CACACGCTGCGCTGCTGCTGCTGCCATGCCATGTTCCCA--- 3512
OY 435 AsnGlnLysaspValTrpLysPheLysGlyLeuasnPheLeuLysCysValProArgPhe 454
Db 3513 -----ATCTGTGCTGGA 3524
OY 455 LysArgArgGlySerHisCysGlyProGlnAlaProSerSerHisProThr 471
Db 3525 CAACCTGTGGAGAACGTTCTTGGAGAGTGGGAGTGTATCATCAACC 3575

```

RESULT 3

US-09-197-636-3

Sequence 3, Application us/09197636

Patent No. 6239267

GENERAL INFORMATION:

APPLICANT: DUCKWORTH, DAVID

APPLICANT: HAYES, PHILIP

APPLICANT: MEADOWS, HELEN

APPLICANT: DAVIS, JOHN

TITLE OF INVENTION: NOVEL COMPOUNDS

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ratner & Prestia

STREET: P.O. Box 980

CITY: Valley Forge

STATE: PA

COUNTRY: US

ZIP: 19482-0980

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,636
FILING DATE: 23-NOV-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9805137.8
FILING DATE: 12-MAR-1998
APPLICATION NUMBER: UK 9815791.0
FILING DATE: 21-JUL-1998
APPLICATION NUMBER: UK 9819278.4
FILING DATE: 03-SEP-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30075
TELECOMMUNICATION INFORMATION:
TELEPHONE: 601-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-197-636-3

Alignment Scores:
Pred. No.: 0.941
Score: 100.00
Percent Similarity: 31.72%
Best Local Similarity: 19.92%
Query Match: 3.40%
Length: 4803
Matches: 103
Conservative: 61
Mismatch: 193
Indels: 160
Gaps: 23

US-09-622-964-29 (1-551) x US-09-197-636-3 (1-4803)

QY 26 GlySerIleTyrLysLeuLeuTyrGlyLupheLeuValPheIlePheLeuTyrTyrSer 45
DB 2292 GGAAGAAGATCCTGCTGTTAGGAGA-----GCTACCTCTTT 2330
QY 46 IleArgLysLeuTyrArgMetValLeuSerSerAspGlnLeuLeuPheGluLysLeu 65
DB 2331 TTCCGAGGATTCAGTATTTCTG-----CAGAGCGCGCGTCGATGAG 2375
QY 66 AlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPheValLeuGlyPheTyr 85
DB 2376 ACCCTGTGTGTCGACGCTACAGTGAAGT-----CTTTCTTCTGCACTGCTTC 2429
QY 86 ValPheLeuValAlaSerArgTyrTrpSerGlnTyrGluAsn----- 99
DB 2430 ATGCTGCCACCGTGTGCTGCTGACCTACACCTCAGAGAGTATGCTTCATGTA 2489
QY 100 -----LeuProTrrProAspArgLeuMetIleGlnValSerSerPheValGlu 115
DB 2490 TTCTCCCTGCGCTTGGGCTGGACCAACATGCTCTAC-----TACACCGCG 2534
QY 116 GlyLysAspGluGlnGlyArgLeuLeuArgArgPheLeuTyrAlaIleLeuGly 135
DB 2535 GGTTCACACAGATGGG-----ATCTATGCGCTCATGATA 2570
QY 136 GlnValLeuIleLeuArgSerIleSerThrSerValTyrLysArgPheProThrLeuHis 155
DB 2571 GAGAGATGATCTTCAGAGACCTGTC-----ATCTATGCGCTCATGATA 2615
QY 156 HisLeuValLeuAlaGlyPheMetThr-----HisGluGluHis 168
DB 2616 ATCGCTCTTCTGCTGGGTTTCCACAGCGGTGTCAGCTGATGTAAGACGGAGAAAT 2675

QY 169 LysGlnLeuGlnLysLeuGlyLeuProHisAsnThrPheTrpValProTrpValTrpPhe 188
DB 2676 GACTCCCTGCGCTGACTCCACGTCGACAG----- 2708
QY 189 AlaAsnLeuSerMetLysAlaTyrLeuGlyGlyArgIleArgAspThrValLeuLeuGln 208
DB 2709 -----TGCGGGGGGCTCCGACGAGGCCCC-----GAT 2738
QY 209 SerLeuMetAsnGlnValLysThrLeuArgThrGlnGlyGlnLeuTyrAla----- 226
DB 2739 AGCTCTTACAC-----AGCTTACTCCACCTCGCTGGAGGCTTCAAGTCCACC 2789
QY 227 -----TyrAspTrpIleSerIleProLeuVal 235
DB 2790 ATCGGCAATGGCGCAGCTGAGTTCACCTAGACAGATATGACTTCAAGGCTGCTCATCACC 2849
QY 236 TyrThrGlnValValThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArg 255
DB 2850 CTGCTGCTGCGCTATGTAATTCACCTACATCTCTGCTCAACATGCTCATCGCCCTC 2909
QY 256 GlnPheLeuAsnProAsnLysAspTyrProGlyHisGlu----- 268
DB 2910 ATGGGTGAGACTGTCTCAACAGATCCGACAGAGACCAAGAACATCTGAACTGCAGAGA 2969
QY 269 -----MetAspLeuValValProValPheThrIleLeuGlnPheLeuPheTyr 284
DB 2970 GCCATCACCATCTCGACACGAGAGAGACCTTCAAGTGCATGAGCAAGGCTTCCGC 3029
QY 285 MetGlyTyrPheLysValAlaGlnLeuIleAsnProPheGlyGluAspAspAsp 304
DB 3030 TCAGGC-----AACTGCTGCGAGGTGGGTACACACCTGATGCAAGAGACGACACCG 3083
QY 305 Phe-----GluThrAsnTrpIle----- 310
DB 3084 TGGTGCTGAGGTGAGAGAGTGAACCTGACACCTGGAACCAACGTCGGCATCATC 3143
QY 311 -----IleAspArgAsnLeuGlnValSerLeuLeuSer 321
DB 3144 AAGCAAGACCGGCGACCTGAGGCGCTCAAGCGCACCTGACCTTCTCCCTCGGTCA 3203
QY 322 -----ValAspGlyMetHisGlnAsnLeuProPheMetGluArgAspMetTyrTrpAsn 339
DB 3204 AGCAGAGTTTCAGCAGACAC-----TGAGAG 3230
QY 340 GluAlaAlaProGlnProProTyrThrAlaAlaSerAlaArgSerArgHisSerPhe 359
DB 3231 AACTTGGCCCTGCTCCCTTTAAGAGAGCAAGTCTGACATGACGCTGCTCAG 3290
QY 360 MetGlySerThr-----PheAsnIleSerLeuLysGluAspLeuGluLeu 375
DB 3291 CCCGAGAGATTTATTCGACAGATTTTCAAGGCTCTTGAAGCCAGAGCGCTGAGGTC 3350
QY 376 TrpSerLysGlnGlnAlaAspThrAspLysGlyGlySerGlyTyrSerSerIleGly 395
DB 3351 TTCAAGAGCTCGCCCTCCCGGAGAGAGTGAAGACGACGACGACACACCTGTC 3407
QY 396 CysPheLeuGlyLeuGlnProLysAsnTyrHisLeuProLeuLysAspLeuLys---Thr 414
DB 3408 -----AACCTGCGGCTTTAGAGACCCCTTCCACGCGGCGCTCGAGGAGACA 3458
QY 415 LysLeuLeuCysSerLysAsnProLeuLeuGlnGlyGlnCysLysAspAlaAsnGlnLys 434
DB 3459 CCAGTCTCTGT---CAGCAGCGCTGCGCTGCTGCTGCTGCGCCACATGTTCCCAA--- 3512
QY 435 AsnGlnLysAspValTrpLysPheLysGlyLeuAspPheLeuLysCysValProArgPhe 454
DB 3513 -----ATCTGCTCGTA 3524
QY 455 LysArgArgGlySerHisCysGlyProGlnAlaProSerSerHisProThr 471
DB 3525 CAAGCTGTGGAAGCGTTCTTGGAAGCATGGGAGATGATGACATCCAAAC 3575

RESULT 4

1893 CCTCGTCCTTCTTCTCTCC---TCTTCATCCCGGTTCAAGTTTAAGTCGACGCCACCC 1943
 QY 372 spleu-----GluEtrpSerLysGluGluAlaAspThrAspLys 387
 Db 1950 CGGTAGAGCCGGCAGCGGGCAGCTGGAGAGACAGCTCCAGAGGACTGAC---AAGA 2006
 QY 387 InsertGlyTrpSerSerThrIleGlyCysPheLeuGlyLeuGlnProLysAsnTrpHisL 407
 Db 2007 GCAGGTGGGCGAGTGGCTCGGGGTGACGCGTGGAGGGGCGAGTGCAGTACTCCCCAC 2066
 QY 407 euProLeuLysAspLeuLysThrLysLeuLysSerLysAsnProLeuGlnGly 427
 Db 2067 CGCCACCAACCCAGATTAAAGTGGAG-----CCCATTCAGAAAGAG 2108
 QY 427 IncysLys-----AspAlaAsnGlnLysAsnGlnLysAspValTrpLys 442
 Db 2109 AGTGGGAGAGGTGGAGGTGACTGACTACATCAAGTCAGAAATGAGAGATGGGAGGTGT 2168
 QY 442 helys-----GlyLeuAsp 447
 Db 2169 TCAGAGCCCCCGTGGCCCGCTGCTCCGCCCAAGCAGAGCCGGAGAGGCCCGGGG 2228
 QY 447 heLeuLysCysValPro-----ArgPheLysArg 457
 Db 2229 TGGCCCGAGTCATGCCCTTAACTCGCTTAAAGCGCGCTGAGTGAAGACTGTGCGCC 2288
 QY 458 -----GlySerHisCysGlyProGlnAlaProSerSerHisProThrGluGlnSerAla 476
 Db 2289 TGGAGGGGGGGGGCTGCTGCTGGGGGCGCTGAAGATAGAGGTGAGGCAAGAGGTGC 2348
 QY 476 roSerSerSerAspThrGlyAsp-----GlyProSerThrAspTrpGlnGluIleCysH 494
 Db 2349 GTGGGAGCGTGGCCCTGGGAGCTGGGGGACCCCTTACCCACGAGCGGGAGACTCTG 2408
 QY 494 IsMetLysLysLysThrValGluPheAsnLeuAsnIleProGlnSer 509
 Db 2409 ACCTTCAGACGCCACAGCCCAACTCTCCCTGGAGACACCGAGATTC 2455

DG
1787 GGAGACATCTGTCTGTGTAGAGA-----GTCATCTTTT 1825

OY
46 ILeArgLLeuTYrArgMeValLeuSerSerAspCnGlnLeuPheGLyLeu 65

DG
1826 TTCGAGGAGATTCACTATTTCCTG-----CAGAGCGCCGGCATGAAG 1870

OY
66 ALaleuTYrCysAspSerTYrIleGlnLeuIleProIleSerPheValLeuGLyPheTYr 85

DG
1871 ACCCTGTTTGAGACACTACAGTAGATG-----CTTTCTTTTGAGCTACTGTTC 1924

OY
86 ValThrLeuValValSerAlaTYrPTTPSPserGlnTyroLysN----- 99

DG
1925 ATGCTGGCCACCGTGGTGCCTACTTCACGCCACCCTAAGAGAGATGAGCTTCATGTA 1984

OY
100 -----LeuProTPProAspArgLeuMetIleGlnValSerPheValGlu 115

DG
1985 TTCTCCCTGCTTGGGCTGGAGCACCAATGCTCTAC-----TACACCCGC 2029

OY
116 GLyAspApGlnGlnLYrArgLeuLeuAlaGlyThrLeuIleArgTYrAlaIleLeuGLy 135

DG
2030 GGTTCACAGATGAGG-----ATCATGCGCTCATGATA 2065

OY
136 GlnValLeuIleLeuArgSerIleSerThrSerValTYrLysArgPheProThrLeuHis 155

DG
2066 GAGAAATATCTCTGAGACCTGAG-----CTTTATATGTTTCTAC 2110

OY
156 HsLeuValLeuAlaGLyPheMethr-----HsGLyGluHis 168

DG
2111 ATCGTCCTTCTGTGGGTITTCACACCGCTGTGACCGTAAATGAGAGGGAAGAT 2170

OY
169 LysGlnLeuGlnLysLeuGLyLeuProHisAnThrPheTrpValProTrpValTrpPhe 188

DG
2171 GACTCCCTGCGCTGAGTCCAGCTCCGACAG----- 2203

OY
189 AlaAsnLeuSerMetLysAlaTYrLeuGLyGlyArgIleArgSphrrValLeuLeuGln 208

DG
2204 -----TGGCGGGGCTGCTGACAGGCCCCCC-----GAT 2233

OY
209 SerLeuMetansGlnValCysThrLeuArgTYrGlnGlyGlnLeuTYrAla----- 226

DG
2234 AGCTCTCAAC-----AGCTTACTCCACCGCTGAGAGCTTTCAAGTTACCC 2284

OY
227 -----TYrAspTriPleserIleProLeuVal 235

DG
2285 ATCGCATGGGCGACCGGAGTTCAGAACATTAAGTCAAGGCTCTTCATCANC 2344

OY
236 TYrThrGlnValValThrValAlaValTYrSerPhePheLeuAlaCysIleuIleGlyArg 255

DG
2345 CTCTGCTGCGCTATGTAATTCACCTACCTACCTCTCTCTCAACATGCTCATGCCCTC 2404

OY
256 GlnPheLeuAsnProAsnLysAspTYrProGLyHisGlu----- 268

DG
2405 ATGGTGAACACTGTCAACAAGATCCGACAGAGAACAGAACATCTGGAAGCTGACAGA 2464

OY
269 -----MetaspLeuValValProValPheThrIleLeuGlnPheLeuPheTYr 284

DG
2465 GCCATCACCATCTGTGACAGGAGAAGAGCTTCCTTAAGTCACTAGAGAGCCCTCCGC 2524

OY
285 MetGLyTriPLeuValValagGlnLeuLeuAsnProPheGLyGluAspAspAsp 304

DG
2525 TCAGGC-----AACCTGCTGAGGTGGGGTACACACCTGATGCAAGGACGACTACCG 2578

OY
305 Phe-----GluThrAsnTriPLe----- 310

DG
2579 TGGTGTTCAGAGGTGAGAGAGTGAATGGACCACTGGAACCAACCAAGCGGCATCATC 2638

OY
311 -----IleasPArgsnLeuGlnValSerLeuLeuSer 321

DG
2639 AACGAGACCGCGGCACTGTGAGGGCGGTCAAGCGCACCTCGAGCTTCTCCGCGGTCA 2698

OY
322 -----ValaspGlyMethHisGlnAsnLeuProPrometGluArgspMethTYrTrpAsn 339

OY 285 MetGlyTrpLeuLysValAlaGlnLeuIleAsnProPheGlyGluAspAsp 304
DB 3030 TCAGGC-----AGCTGCTGCAGGTGGGTACACACCTGATGGCAAGGACACTACCGG 3083
OY 305 Phe-----GluThrAsnTrpIle----- 310
DB 3084 TGGTGCCTCAGGTGCAGCGAGGTGAACTGACACCTGGACACCAACGTTGGGCATCATC 3143
OY 311 -----IleAspArgAsnLeuGlnValSerLeuLeuSer 321
DB 3144 AACGAGACCCGGGCAACTGTGAGGCGTCACAGCCGACCTGACTCTCCCTGGCGTCA 3203
OY 322 -----ValAspGlyMetHisGlnAsnLeuProMetGluArgAspMetGlyTrpAsn 339
DB 3204 AACAGAGCTTCAGCGAGACAC-----TGGAAG 3230
OY 340 GUAAlaAlaProGlnProProTyrThrAlaAlaSerAlaArgSerArgArgHisSerPhe 359
DB 3231 AACCTTGGCCCTGCTCCCTTTAGAGAGCGCAAGTCTCGAGATAGGCACTGCTCAG 3290
OY 360 MetGlySerThr-----PheAsnIleSerLeuLysGlyAspLeuGluLeu 375
DB 3231 CCCGAGGAAGTTTATCTCGCAGACAGTTTTCAGGCTCTCGAAGCCAGAGAGAGCTGAGCTC 3350
OY 376 TrpSerLysGluGluAlaAspThrAspLysGlySerGlyTrpSerThrIle 394
DB 3351 TTCAGAGAGTCTGCCGCGTCCGGGGAGAGAGAGAGCTCAGCAGACAGCAGCTGTC 3407

RESULT 8
US-09-188-930-25
Sequence 25 Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Murison, Rene
TITLE OF INVENTION: Compositions Isolated From Skin Cells
TITLE OF INVENTION: and Methods For Their Use
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 1230
TYPE: DNA
ORGANISM: Rat
US-09-188-930-25

Alignment Scores:
Pred. No.: 0.191
Score: 97.00 Length: 1230
Percent Similarity: 35.11% Matches: 84
Best Local Similarity: 22.34% Mismatches: 48
Query Match: 3.29% Indels: 121
DB: 3 Gaps: 19

US-09-622-964-29 (1-551) x US-09-188-930-25 (1-1230)

OY 14 LeuGlySerPheSerLeuLeuLysTrpArgGlySerIleTyrLysLeuLeuTyr 33
DB 363 TTGGGCTACAGCAGTTCACACCCCTGGGATGAGAGCTTCACGATACAGCA 421
OY 34 GlyGlu-----PheLeuValPheIlePheLeuTyr---TyrSerIleArg 47
DB 422 GGCAGGTGCTGTCTGTACTTCTGTAGTGTGTTCTCTGCTGCTCAGTACGCTTCGG 481
OY 48 GlyLeuTyrArgMetValIleSerSerAspGlnGlnLeuLeuPheGluLysLeuAlaLeu 67
DB 482 AGACTCTTCGAGAGTCTCTAC-----GTCACGCTC 511

OY 68 TyrCysAspSerTyrIleGlnLeuIleProIleSerPheValIleGlyPheTyrValThr 87
DB 512 TTCCTACACAGCCCATTCACGCTGTGAGTACTGTTCGGCTGCTGCTACTATGCTCTT 571
OY 88 LeuValValSerArgTyrTrpSerGlnTyrGlnAsnLeuProTrpProAspArgLeuMet 107
DB 572 GTTGGCCCTGACC----- 583
OY 108 IleGlnValSerSerPheValGlnGlyLysAspGlnGluArgLeuLeuArgArgThr 127
DB 584 -----GTCAGCACCAGAGTCCCGCATGAAAGACAGAGAGTACGCTCTGGGAGAGAT 637
OY 128 LeuIleArgTyrAla-----IleLeuGlnValLeuIleLeuArgSerIle 143
DB 638 CTACTGCTACAGCTGCTGCTGCTCCACATCTTGGAGATGATGATGTTCTTCTGCTCC 694
OY 144 SerThrSerValTyrLysArgPheProThrLeuHisIleValLeuAla----- 160
DB 695 TCTGCCCATCAGTATAG-----TCCACGCTCATCTCGAATCTCAGG 739
OY 161 -----GlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLeu 176
DB 740 AGAATATGAGAAAGTGTGCTGATCCACCTCCACAGACAG----- 778
OY 177 ProHisAsnThrPheTrpValProTrp---ValTrpPheAlaAsnLeuSerMetLysAla 195
DB 779 -----ATCCCTTTGGAGAGCTGTTGAGTATGATGCTCTCTGCTAAC 820
OY 196 TyrLeuGlyLysArgIleArgAspThrValLeuLeuGlnSerLeuMetAsnGluValLys 215
DB 821 TACCTA----- 826
OY 216 ThrLeuArgThrGlnCysGlyLeuLeuTyrAlaTyrAspTrpIleSerIlePro----- 233
DB 827 -----GCAGAGCTGATGATATCATCTGCATCTGCATCTGCTACCTTCGGG 868
OY 234 -----LeuValTyrThrGlnValValThrValAlaValIleTyrSerPhePheLeuAla 250
DB 869 CTCACAGACGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
OY 251 CysLeuIleGlyArgGlnPheLeuAsnProAsn---LysAspTyrProGlyHisGlnMet 269
DB 929 GCGTCTTCACACAGAGTCTACAAAGCATTGTGCTCTCCAAAGCATTAGGAAA 988
OY 270 AspLeuValValProValPheThrIleLeuGlnPheLeuPheTyrMetGlyTrpLeu--- 288
DB 989 -----GCTTCTCTCCCGCTTCTTGTGTAACAGGCTTATGCTG 1027
OY 289 LysValAlaGlnGlnLeuIleAsnProPheGlyGluAspAspAspPheGluThrAsn 308
DB 1028 AAGAGCCGACCCAGGAGTAC-----GCTTCTCTCTCTGAGAGCGCTGAGACAGGC 1078
OY 309 TrpIleIleAspArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGln--- 327
DB 1079 -----TGAAATACACTTCTGACAGCTGCGCGCCGCGCCAGGCT 1114
OY 328 -----AsnLeuPro-----ProMetGluArgAspMetGlyTrpAsn 339
DB 1115 GTCACGAGAGTGGCGCTCTTCACAGACCTCATCTCGAGAGACCAAGCAGATCCGG 1174
OY 340 GUAAlaAlaProGlnProProTyrThrAlaAlaSerAlaArgSerArg 355
DB 1175 CTCGCGCGCCCGCGCTGTCAGCCCGGGGATTCACATGATTCAGA 1222

RESULT 9
US-09-188-930-201
Sequence 201 Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew


```

Db      1481 TTCCGAGGANTTCAGTATTTCCTG-----CAGAGGGGGCCGTCGATGAAG 1525
Oy      66 AlaLeuTyrCysaspSerTyrIleGlnLeuIleProIleSerPheValLeuGlyLeuTyr 85
Db      1526 ACCCTGTTTGTGGACAGCTACAGTAGAAG-----CTTTCTTCTTCGAGTACTGTTTC 1579
Oy      86 ValThrLeuValValSerArgTyrTrpSerGlnTyrGluAsn-----99
Db      1580 ATCCTGGCCACCCTGCTGCTGCTACCTTCACCCACCTCAAGAGAGTATGCTGCATGTA 1639
Oy      100 -----LeuProTrpProAspArgLeuMetIleGlnValSerSerPheValGlu 115
Db      1640 TTCTCCCTGGCTTGGCGTGGACCAACATGCTTAC-----TACACCCGC 1684
Oy      116 GlyLysAspGluGluGlyArgLeuLeuArgTyrIleValSerPheValGlu 135
Db      1685 GGTTCACAGCAAGATGGC-----ACCTATGCCGTCATGATA 1720
Oy      136 GlnValLeuIleLeuArgSerIleSerThrSerValTyrLysArgPheProThrLeuHis 155
Db      1721 GAGAGATGATCTCGAGAGACCTGTC-----CCTTCATGTTTGTCTAC 1765
Oy      156 HisLeuValLeuAlaGlyPheMetThr-----HisGlyGluHis 168
Db      1766 GTCGCTCTTCTTGTTCGGGGTTCACACGCGTGGTGCAGCTGATGTAAGCGGGAAGAAAT 1825
Oy      169 LysGlnLeuGlnLysLeuGlyLeuProHisAsnThrPheTrpValProTrpValTyrPhe 188
Db      1826 GACTCCCTCCCTGCTGATCCAGTCCGACAGC-----1858
Oy      189 AlaAsnLeuSerMetLysAlaTyrLeuGlyGlyArgIleArgAspThrValLeuLeuGln 208
Db      1859 -----TGGCGGGGGCCGCTGAGGCCCC-----GAT 1888
Oy      209 SerLeuMetAsnGlyValCysThrLeuArgThrGlnCysGlyGlnLeuTyrAla-----226
Db      1889 AGCTCTACACAC-----AGCTGTACTCCTCCTGAGCTGTCCTCAAGTTTACC 1939
Oy      227 -----TyrAspTrpIleSerIleProLeuVal 235
Db      1940 ATCGGCATGGCGGACCTGAGTTCACCTGAGAACATGATCTCAAGGCTGTTTCATCATC 1999
Oy      236 TyrThrGlnValValThrValAlaValValSerPhePheLeuAlaCysLeuIleGlyArg 255
Db      2000 CTGCTGCTGGCCTAGTAAATTCATCTCCTCATCTCCTGCTCAACATGCTCATGCGCCCTC 2059
Oy      256 GlnPheLeuAsnProAsnLysAspTyrProGlyHisGlu-----268
Db      2060 ATGGGTGAGCTGTCAACACAGATGCGACAGAGAGCAAAACATCTGGAACTGCAGAGA 2119
Oy      269 -----MetAspLeuValValProValPheThrIleLeuGlnPheLeuPheTyr 284
Db      2120 GCCATCACACATCCCTGGACAGAGAGAGAGAGCTTCTTAAGTCAAGAGAGGCTTCCGC 2179
Oy      285 MetGlyTrpLeuLysValAlaGlnLeuIleAsnProPheGlyLysAspAspAsp 304
Db      2180 TCAGGC-----AAGCTGCTGAGGTGGGTACACACCTGATGGCAAGAGCAGCTACCGG 2233
Oy      305 Phe-----GluThrAsnTrpIle-----310
Db      2234 TGGTGGTTCAGGTTGAGAGAGGTGAAGTGCACACCTGGAACACCACTGGGCATCATC 2293
Oy      311 -----IleAspArgAsnLeuGlnValSerLeuLeuSer 321
Db      2294 AACGAGACCCGGGCAACTGTGAGGGCGTCAAGCCGACCTTCTCCCTGGGTCAC 2353
Oy      322 -----ValAspGluMetHisGlnAsnLeuProMetGluArgAspMetTyrTrpAsn 339
Db      2354 AGCAGAGTTTCAGGAGACAC-----TGGAG 2380
Oy      340 GlnAlaValProGlnProTyrThrAlaIleSerAlaArgSerArgGHisSerPhe 359
Db      2381 AACTTGGCCCTGCTGCTCTTTAAGAGAGGCAAGAGTCTCGATAGCAGTGTCTCAG 2440

```

```

Oy      360 MetGlySerThr-----PheAsnIleSerLeuLysGlnAspLeuGluLeu 375
Db      2441 CCCGAGAGAGTTTATCTGCGACAGATTTTCAGGGCTCTCTGAAGCAGAGACGCTGAGTTC 2500
Oy      376 TrpSerLysGlnGluAlaAspThrAspLys 385
Db      2501 TTCAGAGTCTCTCCGCTTCGCGGAGAG 2530

RESULT 11
US-07-717-331F-4
Sequence 4, Application US/07717331F
Patent No. 5484905
GENERAL INFORMATION:
APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
APPLICANT: Stein
TITLE OF INVENTION: A Receptor Protein Kinase Gene
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: Yahwak & Associates
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07717,331F
FILING DATE: June 19th 1991
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2749 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-07-717-331F-4

Alignment Scores:
Pred. No.: 1-2
Score: 95.50
Percent Similarity: 34.11%
Best Local Similarity: 21.68%
Query Match: 3.24%
Matches: 103
Conservative: 59
Mismatch: 178
Gaps: 135
Dels: 27

US-09-622-964-29 (1-551) x US-07-717-331F-4 (1-2749)
Oy      16 SerPheSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 35
Db      37 ACCCTGCTGCTTCTGCTAGTCT-----57
Oy      36 PheLeuValPheIlePheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSer 55
Db      58 TTCTGCTGTTGAT-----CTATTTCATCTGCTGCTTGCATCTATGCTCAACACTTATAC 114
Oy      56 SerAspGlnGlnLeu-----LeuPheGlnLysLeuAlaLeuTyrCysAspSerTyrIleGln 74
Db      115 TCTTGAGAGTCTCTCAATCTCGAGCAATAGAACACTTGTATCTCCCGGTGAGTCT--- 171
Oy      75 LeuIleProIleSerPheValLeuGlyPheTyrValThrLeuValValSerArgTrp--- 93

```

```

Db 172 -----TTCCAGCTTGGTTCTTCAACCTTGGAGACGCTCGGAGTGTAT 216
Qy 94 -----TTPSerGlnTyrGlnAsnLeuProTTPProAspArgLeuMetIleGlnVal 110
Db 217 CTGGGAATATGG-----TATATAAAAGCCCTGG-----AAAACC 252
Qy 111 SerSerPheValGlnGlyLysAspGlnGlyAlaGlyLeuAlaArgGlyLeuIleArg 130
Db 253 TACCGATGGGTCCGACAGACAGAACACCT-----CTCTCCAGTTCTATATGGAACC 303
Qy 131 TyrAlaIleLeuGlyGlnValLeuIleLeuArgSerIleSerThrSerValTyr----- 148
Db 304 CTCAAATCTCTGGCAACATCTGTCTCTAGTACAGTCTACTACCTACTGTTGGTCG 363
Qy 149 -----LysArgPheProThrIleuHisIleValLeuAlaGly 161
Db 364 ACAATCTTACTAGAGAAATGCGAGATCTCCGGTGTACAGAGCTTCTCCACAGGT 423
Qy 162 -----PheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuProHisAsn 179
Db 424 AATTTGTATATAGACATCCACACAAA-----GACTCAAGTGGATTTCTGTGGCAG 477
Qy 180 ThrPheTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGly 199
Db 478 AGTTTCGATTTTCG----- 492
Qy 200 ArgIleArgAspThrValLeuLeuGln-----SerLeuMetAsnGlnValCysThrLeuArg 218
Db 493 -----ACAGATCTTACTTCTCGGAGATGAACACTAGTTACAGATCTCAAAACAGCGGC 546
Qy 219 ThrGlnCysGlyGlnLeuTyrAlaTyrAspTyr-----IleSerIle 232
Db 547 AACAGG-----TTCCTTACATCGTGAAAGTTCAGATGATCCGCAACGCGG 594
Qy 233 ProLeuValTyrThrGlnValAlaThrValAlaValTyrSerPhePheLeuAlaCysLeu 252
Db 595 AATTCGTGTCAAACTCGACATTCGAAAGGGATTCCTGACTTATCTATA----- 648
Qy 253 IleGlyArgGlnPheLeuAsn-----ProAsnLysAspTyrProGlyHisGly 268
Db 649 -----AATCATTTTTTGATTCACACAGCTTGAACGCAAGACGCGCTTGGAATGA 702
Qy 269 MetAspLeuVal-----ValProValPheThrIleLeuGlnPheLeuPheTyrMetGlyTyr 287
Db 703 ATGGAGTTTAACTGGCATCCGAGAGTGCAGGATTAATTACATGCTTAC-----AATTAT 759
Qy 288 LeuLysValAlaGlnGlnLeuIleAsnProPheGlyLysAspAspPheGlnTyr 307
Db 760 ACGGAGACAGTGAAGACATGCTTACTCTG-----CATATGACC 801
Qy 308 AsnThrIleIleAspArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGln 327
Db 802 AACCAAAAGCATCTACTCCAGATTCAGTACAGTGCAGTGCATGATGATGAGCTGG 861
Qy 328 AsnLeuProPheMetGlnArgAspMetTyrTyrAsnGlnAlaAlaPro-----GlnPro 345
Db 862 ATCCCGCATCAGCGGATTCGACCTCTGTGGACTTACCAACGAGCATGTCGATCCG 921
Qy 346 ProGlyThrAlaAlaSerAlaArgSerArgGlnHisSerPheMetGlySerThrPheAsn 365
Db 922 CTTTACTTATGTGACT-----TATCTTACTGT----- 951
Qy 366 IleSerLeuLysLysGlnAspLeuGlnLeuTyrPheLysGlnGlnAlaAspThrAspLys 385
Db 952 -----GACCTAATATACG----- 963
Qy 386 LysGlnSerGlyTyrSerSerThrIleGlyCysPheLeuGlyLeuGlnProLysAsn----- 404
Db 964 -----TCACCTAACTGTAACTGATTAAGGGTTCGTTCCCAAGAACCCG 1008
Qy 405 TyrHisLeuProLeuLysAsp-----LeuLysThrLysLeuLysCys 418

```

```

Db 1009 CACAGCTGGGACTTGAGACGGAACGAGGGGTGTGTAGACAGCAGCATGAGTGT 1068
Qy 419 SerLysAsnProLeuLeuGlnGlyGlnCysLysAspAlaAsnGlnLysAsnGly 438
Db 1069 AATGGAGATGGGTTTTTG-----CGGCTAAACATATGATATTTGGCGGATACATAAGAC 1122
Qy 439 ValTyrPheLysGlyLeuAspPheLeuLysCysValProArg 453
Db 1123 GCAGCTGTGATCGGACATATGATGTGAAAAATGTCAAGAGAGG 1167

RESULT 12
5386025-7/c
; Patent No. 5386025
; APPLICANT: JAT, SCOTT D., ELLIS, STEVEN B., HAROLD, MICHAEL
; M.: CAMPBELL, KEVIN P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/482,384
; FILING DATE: 20-FEB-1990
; SEQ ID NO: 7:
; LENGTH: 3802
5386025-7

Alignment Scores:
Pred. NO.: 2.82 Length: 3802
Score: 94.50 Matches: 83
Percent Similarity: 38.60% Conservative: 66
Best Local Similarity: 21.50% Mismatches: 138
Query Match: 3.21% Indels: 101
Gaps: 18

US-09-622-964-29 (1-551) x 5386025-7 (1-3802)

Qy 25 ArgGlySerIleTyrLysLeuLeu-----TyrGlyGlnPheLeuVal 38
Db 1508 AGGCTCTGTGTGATTAATGTTGGCAACTGAGATGTGAATACAGTACTTCTTGTG 1449
Qy 39 PheIlePhe-----LeuTyrLysSerIleArgLysLeuTyr 50
Db 1448 TTTATGTATTTGGCAATATCTCTGCGCTCTCTTCTCTCGTCCGTCGAACACAT 1389
Qy 51 ArgMetValIleSerSerAsp-----GlnIleLeuPheGlyLys 64
Db 1388 GATTAATCTTATGCGAGTGGCTCGATACATTAATTAAGCAGTGCCTCAAAAGCAA 1329
Qy 65 -----LeuAlaLeuTyrCysAspSer-TyrIleGlnLeuIleProIleSerPhe 81
Db 1328 ACTAAAGCCCTTCTTAATATCTGATTCCTTTTGGTGTGATTTA-----TTCAC 1278
Qy 81 IleuGlyPheTyrValThrLeuValValSerArgTyrPheSerGlnTyrGlnAsnLeu 101
Db 1277 TGCATCTTCAACACTTCTTAATTTCTTACATTTGCTTGACAAAGTGCCTGAATAC 1218
Qy 101 OTTPProAspArgLeuMetIleGlnValSerArgPheValGlnGlyLysAspGlnGly 121
Db 1217 TACATCTGTAGCAATGCTGTTAATAGACCAAGCTTCAACAAATCATCATGAGAGGT 1158
Qy 121 YargLeuLeuArgArgThrLeuIleArgTyrAlaIleLeuGlyGlnValLeuIleLeu 140
Db 1157 TTCCAAATTTTGGAGACGAGATTCGCGATGTGATGATGATGATCCATACAGCTTCC 1098
Qy 141 -----ArgSerIleSerThrSerValTyrLysArgPheProThrLeuHisHis 156
Db 1097 ACTACATCCACACAGATATGATTTTACGGGATCCACACCTTGATGATGATGATG 1038
Qy 157 -----LeuValIleuAlaGly-----PheMetThrHisGly 166
Db 1037 TCTTTCGGTACATATAAAGATCAATCTTGTGGGGTTCGGCTATATACACCATGG 978
Qy 166 YcLHis-----LysGlnLeuGlnLysLeuGlyLeuPro-----HisAsnThrPhe 181
Db 977 AGAAGCTGGTAAATACCGGGCCAGCCAGTGGCATGCAAAACACTGCTCCACAAACAGTGA 918

```

QY 181 etrpValProtrpValtrpPheValAsnLeuSerMetLysAlaTyrLeuGlyIYargI 201
 Db 917 AGGCTCTCTCTCGATTTTGTGAAACGTCATCATAGGACACTGTCCAGTTC 858
 QY 201 eAgaApThValLeuLeuGlnSerLeuMetAsn----- 212
 Db 857 GTTTAACAGCATGTTCATCTTCATAGATGTCAGGATGAGACAGCTGCGTGC 798
 QY 213 ----GluValCysThrLeuAlaGlyThrGlnCysGlyGlnLeuTyrAlaTyrAsp----- 228
 Db 797 ATAGCATCTCTCTCTAAAGTATGATCATCATGAA--AACAGTTGATCCTGCGC 739
 QY 229 ----TrpLeuSerIleProLeuValTyrThrGlnValAlaThrValAlaValTyrSerPh 247
 Db 738 TGCCGTGTTCACTGTCATTTTTCAGAGTCAAGATCATCTCTCTGTATAGTACAA 679
 QY 247 ePheLeuAlaCysLeuIleGlyIYargGlnPheLeuAsnProAsnLysAspTyrProGlyH 267
 Db 678 CTTCATGTCTGCG-----AAAATCTTCCT----- 654
 QY 267 sGluMetAspLeuValAlaProValPheThrIleLeuGlnPheLeuPheTyr----- 284
 Db 653 ----CATTTGGGCTGCTGCACTTCACTTCTCTGCTCCAAAGCCA 613
 QY 285 ----MetGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGlnAspAs 302
 Db 612 GCGGCACCAAGGCTTAGATCTGTCTGAGAACCTTCATGTCTGCTG-CAAAT 554
 QY 302 PAaSPAPheGluThrAsnTrpIleLeuAspArg-----AsnLeuGlnValSer----- 318
 Db 553 TCCACACAGCTGAGCTCATTTTGGTCCACAGATACAAATCTGTATATTTCATANA 494
 QY 319 -LeuLeuSerValAspGlyMetHis-----G 327
 Db 493 ATATCAACAAGCTGATGACTCCACTGTCTTTTGCAGTGTGACAGGCTCTTCTGCG 434
 QY 327 nasnLeuProPmet 332
 Db 433 ATCTATCCACCACATG 418

RESULT 13
 US-08-961-527-5/c
 Sequence 5, Application US/08961527
 Patent No. 6420135

GENERAL INFORMATION:
 APPLICANT: Charles Kunsch
 TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 NUMBER OF SEQUENCES: 391
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

OPERATING SYSTEM: HP Vectra 486/33

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

Prior APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 13926 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-5

Alignment Scores:

Pred. No.:	Score:	Length:
32.4	94.00	13926
Percent Similarity:	33.96%	110
Best Local Similarity:	20.64%	71
Query Match:	3.19%	Mismatches: 167
DB:	4	Indels: 185
		Gaps: 28

US-09-622-964-29 (1-551) x US-08-961-527-5 (1-13926)

QY 7 AsnLysValAlaAsnAlaArgLeuGlySer-----PheSerSerLeu--- 20
 Db 13923 AATAAATCTATCAATAAATGATAGAGAACCTTATAGGTCATATTCTTTTCTTA 13864
 QY 21 LeuLeuGlySTrPArgGlySerIleTyrLysLeuLeuTyrGlyGlnPheLeuValPheIle 40
 Db 13863 GTGCTCTTTTGGCGGCA-----TTCTATCTCTTTCTCTCTGTATCT 13819
 QY 41 PheLeu---TyrTyrSerIleArgGlyLeuTyr-----ArgMetValLeuSer--- 55
 Db 13818 TACCTTTGTACCGGTATACAGAGAGCTTACTAGATGATAGATGATGCTCAGCCT 13759
 QY 56 -----SerAspGlnGlnLeuPheGluLysLeu-AlaLeuTyrCysAs 70
 Db 13758 CTAGCGGCAAAATTTTCATCTCAGTTACTAGTTATTACGAATCTGCGCTTATACGAG 13699
 QY 70 pSerTyrIleGlnLeu-IleProIleSerPheValLeuGlyPheTyrValThrLeuVal 90
 Db 13698 GAATTTCCATTATATTATTATTGTTGGCTATGAGCTGTTTATAGCGATTTGCTTT 13639
 QY 90 aISerArgTrrPrrSerGlnTyrGluAsnLeuProTrrProAspArgLeuMetIleGlnV 110
 Db 13638 ATGCTTCTGCTTC---CAATAT----- 13619
 QY 110 aISerSerPheValGlnGlyLysAspGlnGlyArgLeuLeuArgThrLeuIleA 130
 Db 13619 ----- 13619
 QY 130 rGTYrAlaIleLeuGlyGlnValLeuIleLeuArgSerIleSerThrServalTyrLysA 150
 Db 13618 -----CTTACTGGGGCAATTTATCTCTTGGTTGGCCAAATGAGATTTCTAC 13567
 QY 150 rGpPheProThrLeuHis-----HisLeuValLeuAlaGlyPheMetThrHisGlyG 167
 Db 13566 ACTTTAAGGGCTTTATTAAGAAAGAGCTCAACAGCAAGAGGGGCAAAATGCTA 13507
 QY 167 lu---HisLysGlnLeuGlnLysLeuGlyLeuProHisAsnThrPheThrPvalProTrrY 186
 Db 13506 AGGCTATAGTACAGCATTTTATTTGGCTTACCTTTAGTTTGGCTTGGAGCCCTTGGC 13447
 QY 186 aITrrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyIYargIleArgAspThrVal 206
 Db 13446 TG-----GGCGCGTTCGCGGCTCTGTTT 13423
 QY 206 euLeuGlnSerLeuMetAsnGluValCysThrLeuArgThrGlnIYargGlnLeuTyrA 226
 Db 13422 TGCCCTTGGCG-----GCTTCAGATGCTTTCAGAG 13393
 QY 226 IaTyrAspTrpIleSerIleProLeuValTyrThrGlnValAlaThrValAlaValTyrS 246
 Db 13392 CTTCGACGAGAGCTGCTCATGTGTGTATACCTGCGCTTGGCGCTA-----C 13342
 QY 246 erPhePheLeuAlaCysLeuIleGlyArg-GlnPheLeuAsnProAsnLysAspTyrPro 265

Db 13341 CATTCCTGCTCTAGCTGACCTGCTAGTATGTTTGAACCAATTCGCAAACTTC--- 13285
Qy 266 GlyHisGluMetAspLeuValProVal-----PheThrIleLeu 279
Db 13284 -----ATCCCATCTCGGAATCTCCAAAGAGGGGCTTTTC 13246
Qy 280 GlnPheLeuPheTyrMetGlyTyrPheLeuValAlaGluLeuIleAsnPropheGly 299
Db 13245 TCATATTGTGATGGGCTTCTTCTGCTC---TGTTTGAATGCTT----- 13204
Qy 300 GluAspAspAspPheGluThrAsnTyrPile-----IleAspArgAsnLeuGlnVal 317
Db 13203 -----CAATTTTAAGTCATATTGTAATTAATGAAGAAATATCATATATG 13156
Qy 318 -----SerLeuLeu----- 320
Db 13155 AAAAATGCAAAACATGTTCTTGGACGAGGTGCTGCTTGTGTTGACGCGCTGTTCA 13096
Qy 321 -----SerValAspGlyMetHisGlnAsnLeuProPheMetGluArgAspMetTyrTrp 338
Db 13095 GGCAGATCCGCTAGTACAGACCAAGAAAGTAATGAATGAACGACGACGACAGCT 13036
Qy 339 AsnGluAlaAlaProGlnProProTyrThrAlaAlaSerAlaArgSerArgArgHisSer 358
Db 13035 AGTAAA-----ACAAAGCGACCTAAAGGAAAGAGTGGCTGAT 12997
Qy 359 PheMet-----GlySerThrPheAsnIleSer-----LeuIleLys 370
Db 12996 TTTGAATTGATGAGGAGTAGAGGCAAGACCTACCTTATCTGATTACAGAGCGCAAGAAA 12937
Qy 371 GluAspLeuGluLeuTyrSerLys-----GluGluAlaAspThr 383
Db 12936 GTCTATCTCAAAATTCCTGGGCTCTGTTGCTCATCTGCTGCTGCTGCTTCCAGATACG 12877
Qy 384 AspLys-----LysGluSerGlyTyrSerSerThrIleGlyCysPheLeuGlyLeuGln 401
Db 12876 GATGAGATTGCTTAAGAAGCGTGCATGACTATGTGCTC----- 12838
Qy 402 ProLysAsnTyrHisLeuProLeuLysAspLeuLysThrLysLeuLeuCysSerLysAsn 421
Db 12837 -----TTGACAGTACGTCA 12823
Qy 422 ProLeuLeuGluGlyGlnCysLysAspAlaAsnGlnLysAsnGlnLysAspValTyrLys 441
Db 12822 CCAGACATAGAGGAGAGCAATCTGAAGCGACTTTAAGAT-----TGG--- 12778
Qy 442 PheLysGlyLeuAspPheLeuLysCysValProArgPheLysArgGlySerHisCys 461
Db 12777 TATTAAGGATGGATTATTAATAATCTCCAGTCTAGTTGACCCATCAGGCAAACTTTTG 12718
Qy 462 GlyProGlnAlaProSerSerHisProThrGln 472
Db 12717 GAAACTTATGCTGCTCCGTTCTTACCCCAACCCAA 12685

RESULT 14
US-08-233-008A-1
Sequence 1, Application US/08233008A
Patent No. 5578480
GENERAL INFORMATION:
APPLICANT: Khandke, Kiran M.
TITLE OF INVENTION: Methods For The Isolation And
TITLE OF INVENTION: Purification Of The Recombinantly Expressed chondroitinase
TITLE OF INVENTION: I and II Enzymes From P. Vulgaris
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470-8426
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,008A
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,885-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3980 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 119..3181
US-08-233-008A-1

Alignment Scores:
Pred. No.: 4.61 Length: 3980
Score: 93.00 Matches: 83
Percent Similarity: 34.40% Conservative: 67
Best Local Similarity: 19.04% Mismatches: 145
Query Match: 3.16% Indels: 21
Gaps: 21

US-09-622-964-29 (1-551) x US-08-233-008A-1 (1-3980)

Qy 91 SerArgTyrTrp-SerGlnTyrGluAsnLeuProTyrProAspArgLeuMetIleGlnVal 110
Db 1012 AGCAATAGTGGAAGCAAGGACGACATCTGATATCAAAATCATATTATTATCA 1071
Qy 110 LserGheValGluGlyLysAspGluGlyArgLeuLeuArgTyrThrIleLeu 130
Db 1072 ACCAGAGATCTTAATCCCAAGAT-----AAACAATATTGATTA 1113
Qy 130 GtYrAlaIleLeuGlyGln-----ValLeuIleLeuArgSerIleSerThr---ValTyr 148
Db 1114 TTATGTTATTATTAGTAATTAACGACATTAATGTTTAATTAATTAAGCCGCTTATGTGCT 1173
Qy 148 rLyArgPheProThr-----LeuHisH 156
Db 1174 GGAAGAAAGTCCCAACAAAGGCGCAACTAAAGCAGATGTACTATTATTAAGCAAGCA 1233
Qy 156 sleuValIleuAlaGlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLe 176
Db 1234 TTATATTGATCAAGAGCTTTGTTAAAGGAGCTGTTAGTACAAACCATCAGTGGGATA 1293
Qy 176 uprOnHisAsnThrPheThrValProTyrValIleThrPheAlaAsnLeuSerMetLysAlaTyr 196
Db 1294 CAGTTCTCGTGGTGTGATATTTCACAGCTATTAAGTGTGATGATGCACTAAAGAAAGCA 1353
Qy 196 rLeuGlyLysArgIleAspAspThrValLeuLeuGlnSerLeu----- 210
Db 1354 CCAACAACTCAAGATTATGATTAATCTGATTAATCTGATTAATTAAGTAAAGTAAAGTAA 1413
Qy 211 -----MetAsnGluValCysThrLeuArgThrGln 220
Db 1414 TGATATGAAGTAAGTGTGATAGCTGATGATTAATTAATTAATTAATTAATTAATTAATTA 1473
Qy 220 nCysGlyGlnLeuTyrAlaTyrAspThrIleSerIleProLeuValTyrThrGlnValVa 240

```

Db 1474 ACATTAGCCTTATTATTACTAGAG-----CCTGATGATCAAAAGCGTATCA 1521
Qy 240 lthValAlaVal1YrSerPhePheLeuAlaCysLeu1-----G1 254
Db 1522 CTTCATTATATCTTACCGCATATATATACAGCGGCATTAACGCAAGTCCGCGGTG 1581
Qy 254 YrGslInPheLeuAsnProAsn-----LysAspYrProGlyH1 267
Db 1582 TAAAGATGTTTACGCCCTGATGATGACAGATGCGACATGACAGCAAGCATCCGGCTA 1641
Qy 267 sglMeLaspLeuValVal1ProVal1PheThr1LeuGlnPheLeuPheTyrMet----- 285
Db 1642 CTCTTTC-----CCAGCCTTAAATAATGCTCTGACCTTATTATTATTATTAG 1689
Qy 286 -----GlyTyrPLeuValAlaGlnLeuLeu1 296
Db 1690 CGATACACCAATTTCAGTGGGTGAAGAGTGGTGAATACCTGATAAAGCAAGCGATGTTG 1749
Qy 296 nProPheGlyGluAspAspAspPheGluThrAsnTyrPLeu1Leu1 316
Db 1750 AGCG-----TGATCTACAGT-----AATCAGA 1773
Qy 316 nValSerLeuLeuSerValAspGlyMetHis-----GlnAsnLeuProPhe 332
Db 1774 AGTTGCA---TTACCGCTTGACAGAGACACCCCTTTTACCTACCTCGTTAAATCAGT 1830
Qy 332 tgluarGaspMetYrTyrP-----AsnGluAlaAlaProGlnProPro----- 346
Db 1831 CGCTCAAGCTATTACTGCTGCTGATGCTGCAATGCTGCAAAATCAGCTGATAAAGCACTGC 1890
Qy 347 -----TyrThrAlaAlaSerAlaArgSerArgRgh1Ser 358
Db 1891 ATCTATTATCTTCTGCGATTAGTGATMAAACAAATGATCAATCAGTATTTTGGAGA 1950
Qy 359 -----PheMetGlySerThrPheAs 365
Db 1951 AACTATTACACAGCGCTTACCTCAAGCTTTTATGCTTATGCTGATGCGCTTTGG 2010
Qy 365 n1le-----SerLeuLysGluAspLeuLeuLeu1 377
Db 2011 TATTCATCCTTGCGCAAGATMAAATGATGACACTGAAGCTTATTAACCAATGTTGTC 2070
Qy 377 rLysGluGluAlaAspThrAspLysGluSerGlyTyrSerSer----- 392
Db 2071 ATCGAAATTTATTAACAAGATTAACCTTATGCGCTTACCAAGATCATGTCGCTCA 2130
Qy 393 -----Thr1leGlyCysPheLeu-----GlyLeuGlnProLysAsnTyr----- 405
Db 2131 AATAGTAGATTAATGCTCGCACCTTACAGGCGCTATCAGCAAGAGTTGGGATTTGAA 2190
Qy 406 -----HisLeuProLeuLysAspLeu 412
Db 2191 TAGATGCAAGGGCAACCACTATTCCTCTCTTAAGACTTA 2236

```

RESULT 15

US-08-233-008A-5
Sequence 5, Application US/08233008A
Patent No. 5578480

GENERAL INFORMATION:

APPLICANT: Khandke, Kiran M.
TITLE OF INVENTION: Methods For The Isolation And
Purification Of The Recombinantly Expressed chondroitinase
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,008A
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,885-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3980 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 188..3181
US-08-233-008A-5

```

Alignment Scores:

```

Pred. No.: 4.61 Length: 3980
Score: 93.00 Matches: 83
Percent Similarity: 34.40% Conservative: 67
Best Local Similarity: 19.04% Mismatches: 145
Query Match: 3.16% Indels: 141
DB: 1 Gaps: 21

```

US-09-622-964-29 (1-551) x US-08-233-008A-5 (1-3980)

```

Qy 91 SerArgTyrPTr-SerGlnTyrGluAsnLeuProTyrProAspArgLeuMet1leGlnVa 110
Db 1012 AGCAATATGTGTGACGCGCAAGCAGACATCTGATCAGTACGTAACAAATCATATTATTA 1071
Qy 110 lSerSerPheVal1GluGlyLysAspGluGluGlyArgLeuLeuArgThrLeu1leAr 130
Db 1072 ACCAGAGATCTTACCTCCACAGT-----AAACAATATTGATTA 1113
Qy 130 gTyrAla1leLeuGlyGln---ValLeu1leLeuArgSer1leSerThrSer---ValTyr 148
Db 1114 TTATGTTATTTTATGTTATTAACAGACATTAATGTTAATTAATTAAGCGCTTATGTC 1173
Qy 148 rLysArgPheProThr-----LeuH1sh1 156
Db 1174 GGAAGAAAGATCCACACCAAAAGGCGCAACTAAAGATGATCTTATTATGACAAGCA 1233
Qy 156 sLeuVal1leAlaGlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLe 176
Db 1234 TTTATTAGATCAAGCGCTTGTTAAGGAGTCTTAGTACGACAAACCATCACTGGGGATA 1293
Qy 176 uProHisAsnThrPheThrVal1ProTyrVal1TrpPheAlaAsnLeuSerMetLysAlaTyr 196
Db 1294 CAGTTCTGCTGTTGATATTTTCCACGTTATTAATGTCGTATGACATAAAGAACCA 1353
Qy 196 rLeuGlyGlyArg1leArgAspThrVal1leLeuLeu1leSerLeu----- 210
Db 1354 CCTCAAACTCAAGTTATGATTCATCTCTGATTCACGATGATTTAAAGTAGTTT 1413
Qy 211 -----MetAsnGluVal1CysThrLeuArgThrG1 220
Db 1414 TGATATGAAGAAAGTGTGATAGCTGATGATCTGATGATTTATTTCAATACCTTATCTGCCCA 1473
Qy 220 nCysGlyGlnLeuTyrAlaTyrAspTrp1leSer1leProLeuVal1YrThrGlnValVa 240
Db 1474 ACATTAGCCTTATTATTACTAGAG-----CCTGATGATCAAAAGCGTATCA 1521

```

OY 240 LrhVAlAlaValTYrSerPheLeuAlaCysLeuIle-----G1 254
Db 1522 CTtAGTTAATACTTACGACCATTTATCATCTGGCGCATTAACGCAAGTGCACCGGCTGG 11
OY 254 YArGgLnPheLeuAsnProAsn-----LysAspTYrProGlyH1 267
Db 1582 TAAAGATGTTTACGCCCTGTAGTACGATGGCGACATGAGCAAGCAACTATCCGGGCTA 1641
OY 267 sGluMetAspLeuValProValPheThrIleLeuGlnPheLeuPheTyrMet-----285
Db 1642 CTCTTTC-----CCAGCCTTAAAAAGCCTCAGCTTATTTATTATTACG 1689
OY 286 -----GlyTrpLeuValAlaGlnLeuIleAs 296
Db 1690 CGATACACCATTTTCAGTGGGTGTAAGATGGTTGGAATACCTGATAAAAAAGCATGGTTTC 1749
OY 296 nProPheGlyGlnAspAspAspPheGlnThrAsnTrpIleIleAspArgAsnLeuG1 316
Db 1750 AGCG-----TGATCTACAGT--AATCAGA 1773
OY 316 nValSerLeuLeuSerValAspGlyMetHis-----GlnAsnLeuProProHe 332
Db 1774 AGTTGGA--TTACCCCTTGCAAGAGACACCTTTTAACCTACCTTCCTTAAATACAGT 1830
OY 332 tGluArgAspMetTYrTrp-----AsnGluAlaLProGlnProPro-----346
Db 1831 CGCTCAAGCTATACTGCTTGCATGTCTGCAAAATCATCTGCCTGATAAAAACACTTGC 1890
OY 347 -----TYrThrAlaAlaSerAlaArgSerArgHisSer-----358
Db 1891 ATCTATTATCTTGGCATTTAGTATAAACACAAATAGATCACTGCTATTTTTGAGA 1950
OY 359 -----PheMetGlySerThrPheAs 365
Db 1951 AACTATTACACCAGCGCTTTTACCTCAAGCTTTCTATGCCCTTTAATGGCGGCTTTGG 2010
OY 365 nIle-----SerLeuLysGlnAspLeuGlnLeuTrpSe 377
Db 2011 TATTATCTGTTGGCAAGATAAATGCTGACACTGAAAGCTTATTAACACCAATGTTTGTGTC 2070
OY 377 tLysGlnGlnAlaAspThrAspLysLysGlnSerGlyTyrSerSer-----392
Db 2071 ATCTGAATTTATTAACAAGATACCCGTATAGCCGCTTACCAAGTCATGCTGTGCTCA 2130
OY 393 -----ThrIleGlyCysPheLeu-----GlyLeuGlnProLysAsnTyr-----405
Db 2131 AATAGTGAGTAAATGCTGCGAGCTTTCACAGAGGCTATACACAGAAAGCTGGGATGGAA 2190
OY 406 -----HisLeuProLeuLysAspLeu 412
Db 2191 TAGAATGCAAGGGGCAACACTATTCACCTTCCCTTAAAGACTTA 2236

Search completed: July 27, 2003, 14:57:53
Job time : 97.042 secs

/Product- "CGICE protein"

W0943695-A1.

02-SEP-1999.

22-FEB-1999; 99WO-US03790.

25-FEB-1998; 98US-0075941.

18-DEC-1998; 98US-0112926.

(MERI) MERCK & CO INC.

(UYUP-) UNIV UPSALA.

Petukhin K, Caskey CT, Metzker M, Madellius C;

WPI, 1999-540560/45.

P-PSDB; MAY29955.

Human and mouse polynucleotides encoding CGICE polypeptides

Claim 2, Fig 8; 67pp; English.

The present sequence represents the mouse CGICE cDNA sequence, which when mutated is responsible for Best's macular dystrophy (BMD). Polynucleotides encoding CGICE are useful for diagnosing whether a patient carries a mutation in the CGICE gene. Normal and mutated CGICE proteins are useful for identifying activators and/or inhibitors of these proteins, in order to treat BMD. The CGICE gene offers a simpler and cheaper method of diagnosing BMD without the need for the presence of the patient. The gene may also be useful to discover the genetic cause of age-related macular dystrophy.

Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 other;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2156-317	1916	551	0	0	0
2944.00					
100.00%					
100.00%					
20					

US-09-622-964-29 (1-551) x AA221229 (1-1916)

QY 1 MethrillethrrThrrAsnLysValAlaAsnAlaArgLeuGlySerPheSerLeu 20
 DB 11 ATGACTATCACCCTACACAAACAAAGTAGCCCAATCCCGCTCGGTCTGCTGCTC 70
 QY 21 LeuLeuGlySTPArgGlySerLeuLysLeuLeuTyrGlyLupheLeuValPheile 40
 DB 71 CTCTGCTGCTGCGAGCAGCAGCTCTACAGCTCTGCTGATGAGAAATTCCTGCTTCATA 130
 QY 41 PheLeuTyrTyrSerLleArgGlyLeuTyrArgMetValLeuSerSerSerPheLeu 60
 DB 131 TTCTCTACTATTCATCCCTGAGCTCTACAGATGCTCTCTCGATATAGACGCTG 190
 QY 61 LeuPheGlyLysLeuAlaLeuTyrCysAspSerTyrLleGlnLeuLleProLleSerPhe 80
 DB 191 TTGTTGAGAGAGCTGCTCTGATCTGCGACCTACATTCACCTCATTCATTCCTC 250
 QY 81 ValLeuGlyPheTyrValThrLeuValLysSerArgTyrPheSerGlnTyrGlnAsnLeu 100
 DB 251 GTTCTGGTTCTATCTTACATTTGGTGGAGCCCTGCTGAGCCAGTACGAGAACTTG 310
 QY 101 ProTyrProAspArgLeuMetLleGlnValSerSerPheValGlnGlyLysAspGlnGlu 120
 DB 311 CCGTGCCCGCAGCCCTCATGATCCAGGTGCTTACCTTCGTGAGGGCAAGAGAGAGAA 370
 QY 121 GlyArgLeuLeuArgTyrThrLeuLleArgTyrAlaLleLeuGlyGlnValLeuLleLeu 140
 DB 371 GGCCGTTTGTCTGCGCGCAGCCTCATCCGCTACCCCATCCGCGCAAGTGCATCTCTG 430

QY 141 ArgSerLleSerThrSerValTyrLysArgPheProThrLleHisLysLeuValLeuAla 160
 DB 431 CGCAGCATCAGCAGCCGCTCTACAGCGCTTTCCACTCTTCACACCTGCTGCTAGCA 490
 QY 161 GlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLeuProHisAsnThr 180
 DB 491 GCTTTATGACCCCAAGGGGAACTATAGCAGTTGAGAAAGTTGGGCTTACCCACACACA 550
 QY 181 PheTyrValProTyrValTyrPheLleAsnLeuSerMetLysAlaTyrLeuGlyLysArg 200
 DB 551 TTCTGGGTGCCCTGGCTGCTGTTGCCACTTGTCAATGAAAGGCTTATCTTGGAGTCA 610
 QY 201 IleArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgThrGln 220
 DB 611 ATCCGGGACACCGCTCTGCTCCAGAGCCTGATGAAAGTGGTGTGCTTCCCTACTACAG 670
 QY 221 CysGlyGlnLeuTyrAlaTyrAspTyrPheLleSerLleProLeuValTyrThrGlnVal 240
 DB 671 TGTGAGACAGCTGTATGCTACAGCTGATAGTATGATCCATGGTGTACACACAGTGTG 730
 QY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuLleGlyArgGlnPheLeuAsnPro 260
 DB 731 ACACTGCGACTATACAGCTTTTCTTCTGATGCTGATGCGAGGAGCTTTCGAAACCA 790
 QY 261 AsnLysAspTyrProGlyHisGlnMetAspLeuValProValPheThrLleLeuGln 280
 DB 791 AACAGAGACTACCCAGCCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 850
 QY 281 PheLeuPheTyrMetLysThrLeuLysValAlaGlnGlnLleLleAsnProPheGlyGlu 300
 DB 851 TTCTTATCTTACATGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 910
 QY 301 AspAspAspAspPheGlnThrAsnThrLleLleAspArgLeuGlnValSerLeuLeu 320
 DB 911 GACGATGATGATTTTGAAGCTACATGATGATGATGATGATGATGATGATGATGATG 970
 QY 321 SerValAspGlyMetHisGlnAsnLeuProMetGlnArgAspMetTyrTyrAsnGlu 340
 DB 971 TCCGTGGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030
 QY 341 AlaAlaProGlnProProTyrThrAlaAlaSerLleArgSerArgGlnHisSerPheMet 360
 DB 1031 GCAGCGCTCAGCGCCCTTACACAGCTCTTCTCCAGGCTCCCGCATTCCTTCATG 1090
 QY 361 GlySerThrPheAsnLleSerLeuLysGlnAspLeuLleTyrPheSerLysGlnGlu 380
 DB 1091 GCTTCACCTTCAACATCAGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG 1150
 QY 381 AlaAspThrAspLysLysGlnSerGlyTyrSerSerThrLleGlyCysPheLeuGlyLeu 400
 DB 1151 GCTGACACGGATTAAGAAAGAGAGCTGATAGACACCATAGGCTGCTTCTTACGACTG 1210
 QY 401 GlnProLysAsnTyrHisLeuProLeuLysAspLeuLysThrLysLeuLeuCysSerLys 420
 DB 1211 CAAACCAAAAGTACATCTTCCCTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1270
 QY 421 AsnProLeuLeuGlnGlnCysLysAspAlaAsnGlnLysAsnGlnLysAspValTyr 440
 DB 1271 AACCCCTCTCAGAGCCAGGTATGAGATGCAACCAAAAGAAAGAAAGAAAGAAAG 1330
 QY 441 LysPheLysGlyLeuAspPheLeuLysCysValProArgPheLysArgArgLysSerHis 460
 DB 1331 AAATTTAAGGCTGTGACTTCTTGAAGTGTTCAGAGTTTAAAGAGAGAGGCTCCCAT 1390
 QY 461 CysGlyProGlnAlaProSerSerHisProThrGlnGlnSerAlaProSerSerSerAsp 480
 DB 1391 TGTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1450
 QY 481 ThrGlyAspGlyProSerThrAspTyrGlnLleCysHisMetLysLysLysThrVal 500
 DB 1451 ACAGGTGATGGCTTCCACAGATTTACCAAGAAATCTGTCTCATGAAAGAAAGAAAG 1510
 QY 501 GluPheAsnLeuAsnLleProGlnSerProThrGlnHisLeuGlnGlnArgLeuAsp 520

DB 1511 GAGTTAACTGACATTCACAGAGCCCAACAGACATCTCAACAGCGCGTTGAC 1570
 QY 521 GlnMetSerThrAsnIleGlnAlaLeuMetLysGluHisAlaGluSerTyrProTyrArg 540
 DB 1571 CAGATGTACACCAATATACAGCGCTCTATGAAGAGCATGACAGACTCTATCCCTACAGG 1630
 QY 541 AspGluAlaGlyThrLysProValLeuTyrGlu 551
 DB 1631 GATGAGCTGGCACCACCAACCTGTCTCTATGAG 1663
 RESULT 2
 AA21227
 ID AA21227 standard; cDNA; 2229 BP.
 AC AA21227;
 XX
 XX 22-NOV-1999 (first entry)
 XX
 DE Human CGICE short form cDNA sequence.
 XX
 KW CGICE: Best's macular dystrophy; mutation; diagnosis; detection;
 KW BMD; age-related macular dystrophy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 105..1862
 FT /tag= a
 FT /product= "CGICE long form protein"
 FT /transl_except= (pos:465..467,aa:Ser)
 XX
 PN W09943695-A1.
 XX
 PD 02-SEP-1999.
 XX
 PF 22-FEB-1999; 99WO-US03790.
 PR 25-FEB-1998; 98US-0075941.
 PR 18-DEC-1998; 98US-0112926.
 XX
 PA (MERI) MERCK & CO INC.
 PA (UIVP-) UNIV UPSALA.
 XX
 PI Petrukhin K, Caskey CT, Metzker M, Madelius C;
 DR WPI; 1999-540560/45.
 DR P-PSDB; AAY29953.
 XX
 PT Human and mouse polynucleotides encoding CGICE polypeptides
 PS Claim 2; Fig 2; 67pp; English.
 XX
 CC The present sequence represents the human CGICE cDNA sequence, which
 CC when mutated is responsible for Best's macular dystrophy (BMD).
 CC Polynucleotides encoding CGICE are useful for diagnosing whether a
 CC patient carries a mutation in the CGICE gene. Normal and mutated
 CC CGICE proteins are useful for identifying activators and/or inhibitors
 CC of these proteins, in order to treat BMD. The CGICE gene offers a
 CC simpler and cheaper method of diagnosing BMD without the need for the
 CC presence of the patient. The gene may also be useful to discovering
 CC the genetic cause of age-related macular dystrophy.
 CC
 XX
 SO Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 other;
 Alignment Scores:
 Pred. No.: 8.07e-202
 Score: 1907.00 Length: 2229
 Percent Similarity: 74.32% Matches: 380
 Best Local Similarity: 64.15% Conservative: 60
 Query Match: 64.78% Mismatches: 102
 Db: 20 Indels: 50
 Gaps: 10

US-09-622-964-29 (1-551) x AA21227 (1-2229)
 QY 1 MetThrIleThrTyrThrAsnLysValAlaAsnAlaArgLeuGlySerPheSerSerLeu 20
 DB 105 ATACCATCATCTTACACAAAGGCGCTAATGCGCGTATGAGCTCTTCGCCGCTG 164
 QY 21 LeuLeuCysTrpArgLysSerIleTyrLysLeuLeuTyrGlyGluPheLeuValPheIle 40
 DB 165 CTGCTGCTGGCGGGGAGCATCTACAAAGCTCTATATGCGGAGCTTCTATCTTCCCTG 224
 QY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGluLeu 60
 DB 225 CTCTGCTTACATCATCGCGCTTATATATAGCGCTGCGCCCTACAGGAGAACACAGCTG 284
 QY 61 LeupheGlyLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIlePheIleSerPhe 80
 DB 285 ATGTTTGAGAAACAGTACTCTGTTATTCGACAGCTACATCCAGCTATCCCATTTCCCTTC 344
 QY 81 ValLeuGlyPheTyrValThrLeuValValSerArgTyrTrpSerGlnTyrGluAsnLeu 100
 DB 345 GTGCTGGCTTCTACGTGACGCTGTGTGACCGCTGTGTGAGAACAGTACAGAACCTG 404
 QY 101 ProTrpProAspArgLeuMetIleGlnValSerSerPheValGlyLysAspGluGlu 120
 DB 405 CCGTGGCCCGACCGCCCTCATGAGCGCTGTGCGGCTTGTGTCGACAAAGGCAAGAGCA 464
 QY 121 GlyArgLeuLeuArgArgThrLeuLeuArgTyrAlaIleLeuGlyGlnValLeuIleLeu 140
 DB 465 GGCCTGGCTGTGGGCGCCAGCTCATCCGCTACGCCCACTGGGCAACGCTCATCTG 524
 QY 141 ArgSerIleSerThrSerValTyrLysArgPheProThrLeuHisLeuValLeuAla 160
 DB 525 CGCAGCGTCAGACCCGACGCTCAAGCGCTTCCCGACGCCAGCGCTGTGTCAGCA 584
 QY 161 GlyPheMetThrHisGlyGluHisLysGlnLeuGlnLysLeuGlyLeuProHisAsnThr 180
 DB 585 GCGTTATATACCTCCGACAGACAAAGAGTGTGAGAAACTGACCTACCAACACATG 644
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrLeuGlyArg 200
 DB 645 TTCTGGTCCCTGTGGTGTGTGTTGTTCCAACTGTGAATGAGCGCTGTGAGAGTGA 704
 QY 201 IleArgAspThrValLeuLeuGlnSerLeuMetAsnGluValCysThrLeuArgThrGln 220
 DB 705 ATCCGGGACCCATCTGCTGCTCCAGAGCTGTGGAACGATGACACCTTCGTAATCAG 764
 QY 221 CysGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValAla 240
 DB 765 TGTGACACCTGTATGCTTACGACTGTGTATGTATCCACATGCTGTATACAGAGTGTG 824
 QY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnPro 260
 DB 825 ACTGGGCGGTGACAGCTTCTCTGACTGTCTAGTTGGCGGAGTTCTGCAACCA 884
 QY 261 AsnLysAspTyrProGlyHisGlnMetAspLeuValValProValPheThrIleLeuGln 280
 DB 885 GCCAAGGCTTACCTGTGCAATGACCTGACCTGTGTGCGCGCTCACAGTTCCTGAG 944
 QY 281 PheLeuPheTyrMetGlyTyrPheLysValAlaGlnGlnLeuLeuAsnProPheGlyGlu 300
 DB 945 TTCTTCTTATATGTTGGCTGCTGACAGTGTGCAAGAGAGCTCATCAACCCCTTGGAGAG 1004
 QY 301 AspAspAspAspPheGluThrAsnTrpIleLeuAspArgAsnLeuGlnValSerLeuLeu 320
 DB 1005 GATGATGATGATTTTGAACCACTGATTTGTGACAGAAATTGAGGTGTCCTGCTG 1064
 QY 321 SerValAspGlyMetHisGlnAsnLeuProPheMetGluArgAspMetTyrTrpAsnGlu 340
 DB 1065 GCTGTGATGATGATGACAGACAGCTGCTCGATGAGACCGGACATGTACGAGATAAG 1124
 QY 341 AlaAlaProGlnProProTyrThrAlaAlaSerAlaArgSerArgArgHisSerPheMet 360
 DB 1125 CCGAGGCGACGCCCCCTTACACAGCTGCTCCGCCAGGTTCCGTCGACCCCTCTTATG 1184

```

OY 361 GlySerThrPheAsnIleSerLeuLysGluAspLeuGluLeuTrpSerLysGluGlu 380
DB 1185 GGCCTCACCTTCACATCAGCCATGACAAAGAGATGAGTCCAGCCCAATCAGAG 1244
OY 381 AlaAspThrAspLysLysGluSerGlyTyrSerSerThrIleGlyCysPheLeuGlu 400
DB 1245 -----GACGAGAGAGATGCTCACCCTGCAATCGCCCTTCCTAGGCTG 1292
OY 401 GlnProLysAsnTyrHisLeuProLeuLysAspLeuLysThrLysLeuLysSerLys 420
DB 1293 CAGTCCATGATCAGTCCATCTCCAGGAGCAACTCAAGACCAACTAGTGTGCCCAAG 1352
OY 421 AsnProLeuLeu-----GluGlyLysCysLysAsp-----AlaAsnGln 433
DB 1353 AGGGATCCCTTCTCCACGAGGCGCTGCCAAAACACACAGCAGCCCAAGAACGTT 1412
OY 434 LysAsnGlnLysAsp-----ValTrpLysPheLysGlyLeuAspPheLeuLysCysVal 451
DB 1413 AGGGCCAGAGAACAAACAGGCGCTGAAAGCTTAAGGCTGAGAGCCCTTCAAGTCTGGC 1472
OY 452 ProArgPheLysArgArgLysSerHisCysGlyProGlnAlaProSerSer----- 468
DB 1473 CCAGTATCAGAGGCGCAGTACACTGCGCCCAAGAGCGCCCTCAGCCCACTGCC 1532
OY 469 -----HisProThrGluLysSerAlaProSerSer-----SerAspThrGly 482
DB 1533 ATGTTCTTCCCTCTAGAACATCAGCGCGCTCAAGGCTTCAACAGTGTCAAGGCAATGAC 1592
OY 482 ----- 482
DB 1593 ACCAAAGACAAAGCTTAAGACTGTGAGTCTGGGCGCAAGAAAGTTTGAATGCTC 1652
OY 483 -----AspGlyProSerThrAspTyrGlnGluLysHisMetLysLysThr 499
DB 1653 TCAGAGAGCGATGGGCGCTTGTATGAGCACCAGAGATGTCATCAAGTGAAGAGAAACT 1712
OY 500 ValGluPheAsnLeu-----AsnIleProGluSerProThrGluHisLeuGlnArgArg 518
DB 1713 GTGGAGTTTACCTGACGAGATATCCAGAGATCCCGAAATCCTCAAGAA-----CCT 1769
OY 519 LeuAspGlnMetSerThrAsnIleGlnAlaLeuMetLysGlnHisAlaGluSerTyr--- 537
DB 1770 TTGGAACAAATCACCACCAACATACACATCACTCAAGATCAATGATCTTATGG 1829
OY 538 -----ProTyrArgAspGluAlaGlyThrLysPro 547
DB 1830 GCCTTGAAAAACAGGAGTGAAGCATCTCTAACCCT 1865

```

RESULT 3
AA221228 standard; cDNA; 2429 BP.

AC AA221228;
XX
XX
XX 22-NOV-1999 (first entry)
DE Human CGICE long form cDNA sequence.
XX
XX CGICE; Best's macular dystrophy; mutation; diagnosis; detection;
XX BMD; age-related macular dystrophy; ss.
OS Homo sapiens.
XX
XX
XX
FH Key Location/Qualifiers
FT CDS 105..1412
FT /tag= a
FT /product= "CGICE short form protein"
XX
XX W09943695-A1.
XX
XX 02-SEP-1999.

```

PF 22-FEB-1999; 99WO-US03790.
XX
XX 25-FEB-1998; 98US-0075941.
PR 18-DEC-1998; 98US-0112926.
XX
XX (MER) MERCK & CO INC.
PA (UUP-) UNIV UPSALA.
XX
PI Petukhin K, Caskey CT, Metzker M, Madellius C;
XX WPI. 1999-540560/45.
DR P-PSDB; AAY29954.
XX
XX Human and mouse polynucleotides encoding CGICE polypeptides
XX
XX Claim 2; Fig 4; 67pp; English.
XX
XX The present sequence represents the human CGICE cDNA sequence, which
XX when mutated is responsible for Best's macular dystrophy (BMD).
XX Polynucleotides encoding CGICE are useful for diagnosing whether a
XX patient carries a mutation in the CGICE gene. Normal and mutated
XX CGICE proteins are useful for identifying activators and/or inhibitors
XX of these proteins, in order to treat BMD. The CGICE gene offers a
XX simpler and cheaper method of diagnosing BMD without the need for the
XX genetic cause of age-related macular dystrophy.
XX
XX Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 other;

```

Alignment Scores:

```

Pred. No.: 4,06e-194 Length: 2429
Score: 1838.50 Matches: 379
Percent Similarity: 66.52% Conservative: 60
Best Local Similarity: 57.42% Mismatches: 103
Query Match: 62.45% Indels: 119
DB: Gaps: 11

```

US-09-622-964-29 (1-551) x AA221228 (1-2429)

```

OY 1 MetThrIleThrTyrThrAsnLysValAlaAsnAlaArgLeuGlySerPheSerSerLeu 20
DB 105 ATGACCATCATCTTACACAGCAAGTGGCTAATCCCGCTTACGCTCTCCGCGCTG 164
OY 21 LeuLeuGlyTrpArgGlySerIleTyrLysLeuLeuTyrGlyLysPheLeuValPheIle 40
DB 165 CTGCTGTGCTGGCGGCGGCGAGCATCTACAGCTGTATATGCGAGTCTTAAATCTCTG 224
OY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGlnLeu 60
DB 225 CTCTGCTACTACATCATCCGCTTATATAGGCTGGCCCTCAGCAAGAACAGAGCTG 284
OY 61 LeuPheGluLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
DB 285 ATGTTTGAGAACTGACTCTGTATGCGACAGTACATCAGCTCATCCCATTTCTCTG 344
OY 81 ValLeuGlyPheTyrValThrLeuValAlaSerArgTrpTrpSerGlnTyrGluAsnLeu 100
DB 345 GTGCTGGCGCTTCTACATGAGCGTGTGAGCCGCTGTGGAAACAGTACAGAACGCTG 404
OY 101 ProTyrProAspArgLeuMetIleGlnValSerSerPheValGluGlyLysAspGlu 120
DB 405 CCGTGGCCGCGACCGCTCATGAGCCTGTGCTGGGCTTCGTGAAGGCAAGAGAGCA 464
OY 121 GlyArgLeuLeuArgArgThrLeuLeuArgTyrAlaIleLeuGlyGlnValLeuIleLeu 140
DB 465 GGGCGGCTCTCGCGGCGCAGCTCATCGCTACAGCAACGCGGCAAGAGAGAGAGCA 524
OY 141 ArgSerIleSerThrSerValTyrLysArgPheProThrLeuHisLeuValLeuAla 160
DB 525 CGCAGCTCAGCAGCGAGCTCACAAGGCGCTTCCCGACGCCCAAGCTGTGTCAAGCA 584
OY 161 GlyPheMetThrHisGlyGlnHisLeuGlnLeuGlnLysLeuGlyLeuProHisAsnThr 180

```

```

Db      585 GCGTTTATGACTCCGGCAGAACACAGCAGTTGGAGAAACGAGCTACACACACATG 644
Oy      181 PheTPValProTPPValTPPheAlaAsnLeuSerMetLysAlaTyrLeuGlyArg 200
Db      645 TTCTGGGGTCCCTGGGTGGTGTGGCCACCTGTCATGAAAGGGCGGTGGAGGTGCA 704
Oy      201 ILAARGAspThrValLeuLeuGlnSerLeuMetCAsnGluValCysThrLeuArgThrGln 220
Db      705 ATCCGGGACCCCTATCCCTCCCTCCAGAGCTGCTGACAGATGATGAAACACTTGGCTACTG 764
Oy      221 CysGlyGlnLeuTyrAlaTyrAspTyrPheSerIleProLeuValTyrThrGlnValVal 240
Db      765 TGTGGACACCTGTATCTCTACAGCAGCTGATGATGATGATGATGATGATGATGATGATG 824
Oy      241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAspPro 260
Db      825 ACTGGCGGTGACAGCTTCTCTGACTGTCTAGTGTGGCGGCACTTCTGAAACCA 884
Oy      261 AsnLysAspTyrProGlyHisGlnMetAspLeuValProValPheThrIleLeuGln 280
Db      885 GCCAAGGCGCTACCTGGCCATGAGCTGACCTCGTTGGCCGCTTCTGACCTGCTGACG 944
Oy      281 PheLeuPheTyrMetGlyTyrLeuLysVal----- 290
Db      945 TTCTCTCTATGTTGGTGGCTGAGTAAAGT-GGGCTCTCCAGGGCCTGCTGGGCTGGAG 1003
Oy      290 ----- 290
Db      1004 GCATGGCCAGAGGGGTATGCGCCAGCAGCTGTTGAGACAGAGATGATGATGATGATG 1063
Oy      290 ----- 290
Db      1064 GAAAGTCTCAAGGGTAGAAAGCAGCAGCGGTGTGGCGCACACTGTATCCAGCTAC 1123
Oy      291 ----- 291
Db      1124 TCCGGAGGCTGAGCGAGAGAAATCGCTGACCCGGGAGCGGAGGTGTGGTGGCAGAG 1183
Oy      293 GlnLeuIleAsnProPheGlyGlnAspAspAspPheGlnTyrAspTyrIleLeuAsp 1183
Db      1184 CAGCTCATCAACCCCTTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1243
Oy      313 ArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGlnAsnLeuProMet 1243
Db      1244 AGGAATTTGACAGTGTCCCTGTGGCTGTGATGATGATGATGATGATGATGATGATG 1303
Oy      333 GluArgAspMetTyrTyrAsnGlnAlaAlaProGlnProTyrTyrAlaAlaSerAla 1303
Db      1304 GAGCGGACATCTACTGGAATTAAGCCGAGCCAGCAGCCGCCCTTACACAGCTGCTCCGCC 1363
Oy      353 ArgSerArgArgHisSerPheMetGlySerThrPheAsnIleSerLeuLysGluAsp 1363
Db      1364 CAGTTCGCTGAGCCTCTTATGAGCTCCACCTCAACATCAAGCCTGAACAAGAGAG 1423
Oy      373 LeuGlnLeuTyrPheLysGlnGluAlaAspThrAspLysLysLeuSerGlyTyrSer 1423
Db      1424 ATGAGATTCCAGCCCAATCAGAG-----GAGAGAGAGATGCTCAGCTCAGCTG 1471
Oy      393 ThrIleGlyCysPheLeuGlyLeuGlnProLysAsnTyrHisLeuProLeuLysAspLeu 1471
Db      1472 ATCATTTGGCGCTTCTAGGCCGTCAGCTCCATGATCCATCTCCAGGAGCAACATCA 1531
Oy      413 LysThrLysLeuLeuLysSerLysAsnProLeuLeu-----GluGlyGlnCysLysAsp 1531
Db      1532 AGGACCCAACTACTGTGGCCCAAGAGGAATCCCTTTCACAGAGAGGCTGCCCAAAAC 1591
Oy      431 -----AlaAsnGlnLysAsnGlnLysAsp-----ValTyrLysPheLys 1591
Db      1592 CACAAGGACAGCAACAGAACTTGTGGGCGCAGGAAGCAACAAGGCTGGAACCTTAAG 1651
Oy      444 GlyLeuAspPheLeuLysCysValProArgPheLysArgGlySerHisCysGlyPro 1651
Db      1652 GCTGTGGAGCGCTTCAACTGTGGCCCACTGTATCAGAGGCGCAGGCTACTACAGTCCCA 1711

```

```

Oy      464 GlnAlaProSerSer-----HisProThrGlnGlnSerAlaProSerSer 478
Db      1712 CAGAGCGCCCTCCACCCCACTCCCATGCTTCTTCCCTCCAGAACCATCAGCGGCTAAAG 1771
Oy      479 -----SerAspThrGly----- 482
Db      1772 CTTACAGTGTCCACAGGCAATAGACACCAAAAGCCTTAAGACTGTGACTTGGG 1831
Oy      483 ----- 483
Db      1832 GCCAAGAAAGTTTGAATTCCTCTCAGAGAGCGAGGCGCTTGTGAGCAGCAGCA 1891
Oy      492 IleCysHisMetLysLysLysThrValGluPheAsnLeu---AsnIleProGluSerPro 510
Db      1892 GTATCTCAAGTGAAGGAGAAAGCTGTGAGGATTAACTGACGAGATATGCCAGATGCC 1951
Oy      511 ThrGlnHisLeuGlnGlnArgTyrLeuAspGlnMetSerThrAsnIleGlnAlaLeuMet 530
Db      1952 GAAATTCACCTCAAGAA-----CCTTGGACAAATCACCACCACTACACACTACACTC 2008
Oy      531 LysGlnHisAlaGlnSerTyr-----ProTyrArgAspGlnAlaGlyThrLysPro 547
Db      2009 AAGATTCACATGATCTTATTTGGGCTTGGAAAAAGGATGAAAGCATTCCTTAACCT 2068
RESULT 4
ABLI0793
ID ABLI0793 standard; cDNA; 2861 BP.
XX
AC ABLI0793;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
OS pharmaceutical; gene; ss.
XX
PN Drosophila melanogaster.
XX
W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB66690.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
PS Claim 1; SEQ ID NO 26861; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABBI7737-ABBI2072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.int/pub/published_pct_sequences.

```



```

DB: 23 Gaps: 4
US-09-622-964-29 (1-551) x ABL12609 (1-1608)
QY 1 MetThrIleThrTyrThrAsnLysValAlaAsn---AlaArgLeuGlySerPheSerSer 19
DB 1 ATGACGTCTCTTACACCCGCTGAGCGGCACATCCACCATTTGGCTGCTTGGAAG 60
QY 20 LeuLeuLeuGlyTyrPargLysSerIleTyrLysLeuLeuTyrGlyGluPheLeuValPhe 39
DB 61 CTTTGATGATGAGCGCGGCAAGTATCTACAAAGATAATATGGGATCTTCTGCAATTC 120
QY 40 IlePheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGln 59
DB 121 CTGCTCTCTCTTCTTCTTCTGATGCTGTATATATGCGATGCTCAGGAGCGTGTATAG 180
QY 60 LeuLeuPheGlyLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSer 79
DB 181 CCTGTCTTTCAGGACATGTGATGATCTGTATGATGATGATGATGATGATGATGATGAT 240
QY 80 PheValLeuGlyPheTyrValTyrLeuValValSerArgTyrPheSerGlnTyrGluAsn 99
DB 241 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 100 LeuProTyrProAspArgLeuMetIleGlnValSerSerPheValGlyLysAspGlu 119
DB 301 GTTCCTTGGCGGATCCGCTGCGGCTTATGATGATGATGATGATGATGATGATGATGAT 360
QY 120 GluGlyArgLeuLeuArgTyrThrLeuIleArgTyrAlaIleLeuGlyGlnValLeuIle 139
DB 361 CATGCTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 140 LeuArgSerIleSerThrSerValTyrLysArgPheProThrLeuIleHisLeuValLeu 159
DB 421 CTCTCGATGATATCCCACTTATTAAGGCGCGCTTCCAACTGATGATGATGATGATGAT 480
QY 160 AlaGlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLeu----- 176
DB 481 GTGGGTTTGTAAACGCCAACGAGCAAAATATTAAGAGCAATGATGATGATGATGATGAT 540
QY 177 ProHisAsnThrPheTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyr 196
DB 541 AAGACCCCAAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 197 LeuGlyArgTyrIleArgAspThrValLeuLeuGlnSerLeuMetLysGlnValCysThr 216
DB 601 AAGCAAGCTGAATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 217 LeuArgThrGlnCysGlyGlnLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyr 236
DB 661 TTCCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 237 ThrGlnValValThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGln 256
DB 721 ACACAGCTGCTACCCCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 257 PheLeuAsnProAsnLysAspTyrProGlyHisGlyMetAspLeuValValProValPhe 276
DB 781 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 277 ThrIleLeuGlnPheLeuPheTyrMetGlyTyrPheLysValAlaGlnLeuIleAsn 296
DB 841 AACACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 297 PropheGlyGluAspAspAspPheGlyThrAsnTyrIleLeuPargAsnLeuGln 316
DB 901 CTTTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 317 ValSerLeuLeuSerValAspGlyMetHisGlnAsnLeuProPheMetGluArgAspMet 336
DB 961 GTGAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 337 TyrTrpAsnGlnIleAlaPro---GlnProProTyrThrAlaIleSerAlaArgSerArg 355

```

```

DB 1021 TATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 356 ArgHisSerPheMetLysSerThrPheAsnIleSerLeuLys 370
DB 1081 CATCCGAGC-----GCATCAACCGCTGATGATGATGATGATGATGATGATGATGAT 1119
RESULT 6
ABLI0792/c
ID ABLI0792 standard; cDNA; 10760 BP.
XX
AC ABLI0792;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26858.
XX
KW Drosophila: developmental biology; cell signalling; insecticide;
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001MO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
DR P-PSDB; ABB6689.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Claim 1: SEQ ID NO 26858; 21pp + Sequence Listing: English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
XX sequences (ABLI5737-ABLI72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 10760 BP; 3232 A; 2079 C; 2173 G; 3276 T; 0 other;
XX
Alignment Scores:
Pred. No.:
Score: 1.82e-80 Length: 10760
Percent Similarity: 829.00 Matches: 194
Best Local Similarity: 48.46% Conservative: 73
Query Match: 35.21% Mismatches: 120
DB: 28.16% Indels: 165
Gaps: 11
US-09-622-964-29 (1-551) x ABLI0792 (1-10760)
QY 21 LeuLeuGlyTyrPargLysSerIleTyrLysLeuLeuTyrGlyGluPheLeuValPhe 40
DB 5539 CTTTTCAGATGCGCAGAGACATTTACAAAGTGGCTTATGCTTCTGCGCTCTG 5480
QY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGln---- 59

```

```

Db      5479  ACATTTACTATGCGATCAACATGATGCTATCGCTTTGGCTCAACCCCGCACAAAAAGAG 5420
Oy      59  -----
Db      5419  TGAATATAAAAAAGCTCTAGAGAGATGATGAGATTTGGTGGCTTACATTAACATA 5360
Oy      59  -----
Db      5359  TTTATGATTTGAATATATATTTTGAATTTATGATATGATTCATTCATTCAGAA 5300
Oy      59  -----
Db      5299  AGTGTGATGGCTCTACTCAATATACATGACATTAATTAATTAACACAGACATA 5240
Oy      59  -----
Db      5239  GAAAGTTTACTCAAAACATCTGAAAATAGCTCTTGACAACTTCACAAATGGTGTATAT 5180
Oy      60  -----
Db      5179  AAAATTAGCTGATGCTACAAAAATACATATATATGATGAGATCTTCTTCACCGA 5120
Oy      62  -----
Db      5119  TCTGTCTCTTTTAAAGCCCTTACTAATATGCAACTTCTTCAGAACCTTTGAGCC 5060
Oy      65  LeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPheValLeuGlyPhe 84
Db      5059  ATTGTTCAGTACTCTGATAGTATACAGAAACCTATACCTCCCTGCTTCCTGCTGGTTTC 5000
Oy      85  TyrValThrLeuValValSerArgTyrPheSerIleGlnLeuIleProIleProAsp 104
Db      4999  TATCTATGATGATGATGAGCCGCTGGTGGATACAGTACACCTCCATCCCTGGCCAGAT 4940
Oy      105  ArgLeuMetIleGlnValSerSerPheValGlnGlyLysAspLeuIleGlnArgLeuLeu 124
Db      4939  CCCATCGCGCTGTTTGTCACTGCAATGTCATGCGCAGGATGACGAGACACGATGAG 4880
Oy      125  ArgArgThrLeuIleArgTyrAlaIleLeuGlnValLeuIleLeuArgSerIleSer 144
Db      4879  AGCGAACAATATGCGAATGTGCTTGGCTGACAGATGCTGCGAATGTTCG 4820
Oy      145  ThrSerValTyrLysArgPheProThrLeuHisHisLeuValLeuAlaGlyPheMetThr 164
Db      4819  CCGAGGTGAAGAAGCTTCCCGCCCTAATATCTGCTGAAGCGGTCTGCTAAT 4760
Oy      165  HisGlyGlnHisLysGlnLeuGlnLysLeu-----GlyLeuProHis---AsnThrPhe 181
Db      4759  GACAAATGAAAAGACCATCATCGACCATGAACAAGCCCTTCCAGACCTTCGAAAGAC 4700
Oy      182  TyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyGlyArgIle 201
Db      4699  TGGCTGCCCATGCTTTGGCTGCGCAGTATATATACAGGCCGCAAGAAAGAGTCCGCAAT 4640
Oy      202  ArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValLysThrLeuArgThrGlnCys 221
Db      4639  CGTGATGATTTTGTCTGAAGCATCATCATGATGACCTTAATAGTTTTCGTGTCAGTGT 4580
Oy      222  GlyGlnLeuTyrAlaTyrAspTyrIleSerIlePheLeuValTyrThrGlnValValThr 241
Db      4579  GGACCTCTCATCAGTACGATACCATTAATGCTGCTGCTGACACCAAGTGTGTACC 4520
Oy      242  ValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsn----- 259
Db      4519  CTGGCGGTATGTCTACTCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
Oy      260  -----ProAsnLysAspTyrTyrGlyHisGlnLeuSerAspValValProValPheThr 277
Db      4459  GTGTGGGCAATACCATACATCTG---AACAGGTGATCTATACCTTCTGATTTTACA 4403
Oy      278  IleLeuGlnPheLeuPheTyrMetGlyTyrLeuLysValAlaGlnLeuLeuIleAsnPro 297

```

```

Db      4402  ACCTGCGAGTCTTCTTCTACATGATGGTGGCTCAAGGTGGCCGAGTGGCTGATTAATCCA 4343
Oy      298  PheGlyGlnAspAspAspPheGlu----- 306
Db      4342  TTTGGCAAGACGATGATGATTTTGAGGTACGATTAACAACATTTGTTGTTCTGCGA 4283
Oy      307  -----ThrAsnTyrIleLeuAspArgAsnLeuGlnValSer 318
Db      4282  AATCAATATAATATATCTCTCAGTCAACTGATGATGGATGCAATCTTCAGTGTCC 4223
Oy      319  LeuLeuSerValAspGlyMetHisGlnAsnLeuProPheMetLeuArgAspMetTyrTrp 338
Db      4222  TATCTGATCTCCACGATGACATGACATTCGAGCTGTAAAGATCAGTACTGG 4163
Oy      339  AsnGlnAlaAlaPro---GlnProTyrTyrAlaAlaSerAlaArgSerArgArgHis 357
Db      4162  GACGAGGTGTTCCCAACGAGCTGCCCTACACAAATAGCTGCCGAAACGATTCGGGAGAT 4103
Oy      358  SerPheMetGlySerThrPheAsnIleSerLeuLysGlu----- 371
Db      4102  CATCCAGACCGTCCACTGCGCAAGATCGAGGTGCCCAAGATCGCGCATGCCATCGACA 4043
Oy      372  -----AspLeuGlnLeuTyr-----SerLysGlnGlu 380
Db      4042  ATGTGCTCCGTCGCGATCATGCA- AATGGTGTGTAATGTAATGCTGATCCCGCCGCC 3984
Oy      381  AlaAspThrAspLysLysGlnSerLysTyr-----SerSerThrIleGlyCys-PheLe 398
Db      3983  TCGCGATGCGAAGACCCCGATACCGGTCAAGGTCAAGTTCAAGTTCGGGTGCGGACAC 3924
Oy      398  uGlyLeuGlnProLysAsnTyrHisLeuPro 408
Db      3923  CGGCGTCCAAACCGTTACTTATGATTTCCC 3893

RESULT 7
ABLI2636/C
ID      ABLI2636 standard; cDNA: 3592 BP.
XX
XX      ABLI2636;
XX
XX      26-MAR-2002 (first entry)
XX
XX      Drosophila melanogaster expressed polynucleotide SEQ ID NO 32390.
XX
XX      Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical; gene; ss.
XX
XX      Drosophila melanogaster.
XX
XX      WO200171042-A2.
XX
XX      27-SEP-2001.
XX
XX      23-MAR-2001; 2001WO-US09231.
XX
XX      23-MAR-2000; 2000US-191637P.
XX
XX      11-JUL-2000; 2000US-0614150.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PMD, Myers EW;
XX
XX      WPI; 2001-656860/75.
XX
XX      P-P-SDB; ABB68533.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
XX      Claim 1; SEQ ID NO 32390; 21pp + Sequence Listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is

```


CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention of
 CC sequences genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC (AB57737-AB572072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pat_sequences.

XX Sequence 3592 BP; 1064 A; 781 C; 774 G; 973 T; 0 other;

Alignment Scores:

Pred. No.:	9.66e-81	Length:	3592
Score:	824.50	Matches:	203
Percent Similarity:	45.77%	Conservative:	100
Best Local Similarity:	30.66%	Mismatches:	184
Query Match:	28.01%	Indels:	177
DB:	23	Gaps:	15

US-09-622-964-29 (1-551) x AB112636 (1-3592)

```

QY 1 MetThrIleThrTyrThrAsnLysValAlaAsnAlaArg---LeuGlySerPheSer 19
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2904 ATGACGTGACCTATACCTACGCTGAGTACCTGAGTGTCTCAGTAATTTCTGGGG 2845
QY 20 LeuLeuLeu-----
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2844 CTGCTGCTCAGGTTGCTACTCCAAATATATATATTTTCCATTTTAAGCTTTAAAG 2785
QY 23 -----CysTyrArgGlySerIleTyrLysLeuLeuTyrGlyLupheLeu 37
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2784 GACACCTATTTCAAAGATGCGCGGAGTATTACAGTTAAATATGTTGACCTTATT 2725
QY 38 ValPheIlePheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerasp 57
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2724 GTATTTTTCGATGTGCTGTTACTGTCGCAATTACATATGCTGTGTGTAAGCAAGAG 2665
QY 58 GluGln-----
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2664 GCGAAAGGCTAGCGTTATTGATTCCTTTAAATAACAAGCTTTGTCACAGGTTT 2605
QY 62 PheGluLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPheVal 81
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2604 TTGGAGGCAATTATGCTATTTGCGAGAAAATGCGTCCCTCATCTCCGTATGCTGCT 2545
QY 82 LeuGlyPheTyrValIleLeuValIleSerArgTyrTyrSerGlnTyrGluLysLeuPro 101
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2544 TTGGGATTTTTCGTGAGAAATCGTATGACCGATGCGTGGACCAATATACCACTTCT 2485
QY 102 TrpProAspArgLeuMetIleGlnValSerSerPheValGlnGlyLysAspGlnGly 121
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2484 TGCGCAGATGGAATTCCTATTTGATATACACAGATATCCATGCGTATGATAGAGCC 2425
QY 122 ArgLeuLeuArgTyrThrLeuIleArgTyrAlaIleLeuGlyGlnValLeuIleLeuArg 141
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2424 AAGGTAATGCGCGAAGCTATCTAGATATGTGCTGCGCAGGTGATATATATAC 2365
QY 142 SerIleSerThrSerValTyrLysArgPheProThrLeuHisIleLeuValLeuAlaGly 161
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2364 ATGATATACACCGCTGTCAGAGCGCGATTTCCACCTACATCAATATATGAGCTGGC 2305
QY 162 PheMetThrHisGlyGlnHisIleGlnLeuGlnLysLeuGlyLeuProHisAsnThrPhe 181
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2304 TTTTGTGTGAGAAATAGAAATATCTTGGCGCATGATGATGAGAGCCGAGAGGAG 2245
QY 182 -----TyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGly 198
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2244 CCAAGACACTGATGCGCATGCTTGGCGCATGATGATGATGATGATGATGATGATGAT 2185
QY 199 GlyArgIleArgAspThrValLeuLeuGlnSerIleMetLysGlnValCysThrLeuArg 218
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2184 AATTAATATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2125
  
```

```

QY 219 ThrGlnCysGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGln 238
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2124 GCGAATGTGGATTTTGTGCTATATACATGATGATGATGATGATGATGATGATGATGAT 2065
QY 239 ValValThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeu 258
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2064 GTGGTCACCGTGGCCACGATACCTCTTCCTGTCAGTGTCCCTCGACAGCAGCGAAT 2005
QY 259 AsnProAsnLysAspTyrProGlnHisGlnMetAspLeuValProValPheThrIle 278
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 2004 GAAAGTCATTCGACGAGAGATGCTAGATCAGAGATGATGATGATGATGATGATGATGAT 1945
QY 279 LeuGlnPheLeuPheTyrMetGlyTyrPheLysValAlaGlnGlnLeuAsnProPhe 298
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1944 CTCGAGTTTCTTCTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1885
QY 299 GlyGlnAspAspAspAspPheGluThr-----
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1884 GGAGAGATGATGACGACTTTCGAGGTGAGTCTTAAGTTCGATTAACATATACATCAAT 1825
QY 307 -----
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1824 TGCTATAACATAGTTTCTACTATATACATATAGCTTAATGCTTTATTACTGACAGCTC 1765
QY 308 AsnTrpIleIleAspArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGln 327
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1764 AACTGATATATCGATAGAAATCTTACCGTTCCTTATGATGATGATGATGATGATGATGAT 1705
QY 328 AsnLeuProPheMetGluArgAspMetTyrTyrAsnGlnAlaPro---GlnProPro 346
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1704 GACATCCCGGAACTGTTAAGATCATGATATTTGGAAGAGGCTTCCCTAATGAATTCGA 1645
QY 347 TyrThrAlaIleSerIleAlaArgSerArgArgHisSerPheMetGlySerThrPheAsnIle 366
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1644 TACGC--TCAGCCAAAATGCGTCAAAATCTCCAGCAGCTCCACGCGCTATATGGAATC 1586
QY 367 SerLeuLysGlnLysPheLeuGlnLeuTyrPheLys-----GluGln 380
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1585 ACATCATAAAGTAAGTATTTATTTGATTTCTCATTAATGAATAGTAATATCCGAT 1526
QY 381 AlaAspThrAspLysGlnSerGlyTyrSerSerThrIleGlyCys----- 396
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1525 TTTTAGCTCCAGAAAGCAGCATGATGCAATGCTCACTTATGATACCCGTTGCGGAAC 1466
QY 397 -----PheLeuGlyLeuGlnProLysAsnTyrHisLeuProLeuLysAsp 411
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1465 AACTCCGACAAACTTCTCTGGG----- 1442
QY 412 LeuLysThrLysLeuLeuCysSerLysAsnProLeuLeuGlnGlyGlnCysLysAspAla 431
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1441 -----AGACCATTCCTCCACTGCTAGCTACTCAGACTCC 1409
QY 432 AsnGln-----LysAsnGlnLysAspValTyr 440
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1408 ATCCAAATGAAGATGGCGCTCCGAGTGAAGCTTTCAGTGAAGATCAATTTGAT----- 1355
QY 441 LysPheLysGlyLeuAspPheLeuLysCysValProArgPheLysArgGlySerHis 460
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1354 -----TTGCCACATGATGAGCAAGCAACACTCCACAC 1325
QY 461 CysGlyProGlnAlaProSer----- 467
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1324 AAGCGGTGCAACACAACTAGGCTTTGATATATCACTTTATTCGCTTAAATTTT 1265
QY 468 -----SerHisProThrGlnGlnSerAlaProSerSerAspThrGlyAspGlyProSer 486
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1264 TGCTCTATATCTTT-----TCA 1247
QY 487 ThrAspTyrGlnGlnIleCysHisMetLysLysLysThrValGlnPheAsnIle 506
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1246 GTTGACTTAAAGATCTAATGAGCAGCTCGTCCG----- 1211
  
```

QY 507 ProgluserProthgluHnleuglnlgnargyleuapglmetserThAsnIle 526
DB 1210 -----GAACGAGGAGGACATCGACACCGATTTGGATATGCAAGTACCAT--- 1160
QY 527 GlnAlaLeuMetLysGlnHnslAlaGluSerTyProTyArgAspGluAlaGlyThrLys 546
DB 1159 -----CTGCGACATGATACCTCCGCTGCTCTTATAC-GATCAAGTCTTCTCTCC 1110
QY 547 Proval 548
DB 1109 CCCATC 1104
RESULT 8
ABLI2608
ID ABLI2608 standard; cDNA; 3785 BP.
AC ABLI2608;
DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 32306.
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX MO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001MO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
XX DR P-PSDB; ABB68505.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX Claim 1; SEQ ID NO 32306; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
XX sequences (ABLI5737-ABLI2072).
XX The sequence data for this patent did not form part of the printed
XX at ffp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 3785 BP; 1064 A; 764 C; 851 G; 1106 T; 0 other;
Alignment Scores:
Pred. No.: 3.97e-75
Score: 774.50 Length: 3785
Percent Similarity: 55.17% Matches: 165
Best Local Similarity: 37.93% Mismatches: 75
Query Match: 26.31% Indels: 128
DB: 26.31% Gaps: 68
US-09-622-964-29 (1-551) x ABLI2608 (1-3785)

QY 1 MethTrIleHnTYrThAsnLysValAlaAsn---AlaArgLeuGlySerPheSerSer 19
DB 1001 ATGACTGCTCTTATACCCGCTGAGGTGGCAGCATGACGCAATTCGGCTCTTGGAG 1060
QY 20 LeuLeuLeuCys----- 23
DB 1061 CTTTGTGATGAG-CTGAGGACATCTCCGATGGCCCATGGCAAGTTACTAAGATCTGA 1119
QY 24 -----TPrArgLysSerIleTyLysLeuLeuTyrgLgLPheLeuValPhe 39
DB 1120 TCATTCGCGAGATGGCCGCAAGATCTACAGATATAGGTGATCTTCTGGCATTC 1179
QY 40 IlePheLeuTyTySerIleArgGlyLeuTyArg-MetValLeuSerSerAsp----- 57
DB 1180 CTGCTGCTTCTTACTTCACTGCTGATATATATGCTATGCTCTCAGGACGTCGATAG 1239
QY 57 ----- 57
DB 1240 CCGTAGATGATCATAGAAATAGAAAGCTTTAATTATTTAACCCCAACCCCTT 1299
QY 58 -GlnGlnLeuLeuPheGlnLysLeuAlaLeuTyCysAspSerTyIleGlnLeuIlePr 77
DB 1300 AACCCATAGTGTTCGAGGACATCTGTGACTGTCATACCTATAGCAACCTGATTC 1359
QY 77 oIleSerPheValLeuGlyPheTyValThrLeuValIleSerArgTyPrSerGlnTy 97
DB 1360 GCTCTCTTCTGACTGATTTATGTTGGGCACTTATCATAGAACCTGTGGAAATCAATA 1419
QY 97 rGlnAsnLeuProTPrProAspArgLeuMetIleGlnAlaSerSerPheValGlnGly 117
DB 1420 TATCAGACTTCTTCCGCGATCCGCTGCGGCTGATGATGATGCTGCTGCGGCA 1479
QY 117 sAspGlnGlnLysArgLeuArgTyThrLeuIleArgTyValIleLeuGlnGln 137
DB 1480 AGATGACATGCTGCTGATGAGACGACATATATGATGATGCTGCTGATGAC 1539
QY 137 IleuLeuLeuArgSerIleSerThrSerValTyTyArgPheProThrLeuHnslsle 157
DB 1540 TATGCTGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
QY 157 uValLeuAlaGlyPheMetThrHnslGlyHnslsleGlnLeuGlnLysLeuGln 176
DB 1600 AATGAGGTGGTGTGCTAAACGCCAAGCAAGCAATATTAAGCAATGATGTGA 1659
QY 177 -----ProHnslHnslPheTyValProTyValTyPrPheAlaAsnLeuSerMet 194
DB 1660 GTTTCCAAAGCACCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1719
QY 194 sAlaTyLeuGlyGlyTyArgIleArgAspThrValLeuLeuGlnSerLeuMetAsnGln 214
DB 1720 GGCTCGAAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779
QY 214 LysThrLeuArgTyHnslCysGlyLysLeuTyValTyArgTyPrIleSerIleProle 234
DB 1780 CAATTAATTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839
QY 234 uValTyTyThr----- 237
DB 1840 GGTCTATACACAGGTAGTTATGACTCTATTCCTAAGCAGATATTAAGATATATTC 1899
QY 238 -GlnValValThrValAlaValTySerPhePheLeuAlaCysLeuIleGlyArgGlnPh 257
DB 1900 TCAGGTGCTGACCCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1959
QY 257 eLeuAsnProAsnLysAspTyTyProGlnHnslGlnMetLysLeuValValProValPheTh 277
DB 1960 GATCGATCGGATATTAAGCATCAATATATATGATGATGATGATGATGATGATGAT 2019
QY 277 rIleLeuGlnPheLeuPheTyTyMetGlyTyPrLeuLysValAlaGlnGlnLeuLeuAsnPr 297
DB 2020 CACTCTGAGTGTTCCTTCATGAGGCTGCTTAAGTGGCCAGCATGATGATGATGAT 2079
QY 297 oPheGlyLysAspAspAspPheGlnTyHnslTyPrIleLeuAspArgAsnLeuGlnVal 317

Db 2080 TTTTGGGAGGACAGACGACTTGAACCTGACCTGATCGATCGCAACTTCCAGT 2139
 QY 317 LserLeuSerValAspGlyMetHisGlnAsnLeuProPomEGLuAArgSPmetry 337
 Db 2140 GAATTAAGTTCGTTGATGATGAATGACATATCATCCGAGTTGGTGGAGGATCAGTA 2199
 QY 337 rTPAsnGluAlaAlaPro---GlnProProTyrThrAlaAlaSerAlaArgSerArgar 356
 Db 2200 TTGGATGAGGTGTTCCCGCTGAGTTCGATATGCGCTGGAGTCCGATAGCGCGAACA 2259
 QY 356 gHisSerPheMetGlySerThrPheAsnIleSerLeuLys 370
 Db 2260 TCCGGAG-----GCATCAACCGCTCGATGGCATACCAAA 2296
 RESULT 9
 ABL12637
 ID ABL12637 standard; CDNA; 1345 BP.
 AC ABL12637;
 XX
 XX
 DE 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32393.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR P-PSDB; ABB68534.
 XX
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 32393; 21pp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB16176-AB16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SO Sequence 1345 BP; 366 A; 322 C; 323 G; 334 T; 0 other;

Alignment Scores:

Pred. No.: 1.24e-74
 Score: 763.50
 Percent Similarity: 51.748
 Best Local Similarity: 35.388
 Query Match: 25.93%
 Db: 23
 Gaps: 9

'US-09-622-964-29 (1-551) x ABL12637 (1-1345)
 QY 68 TyrcysAspSerTyrIleGlnLeuIleProIleSerPheValLeuGlyPheTyrValThr 87
 Db 7. TATTCGCAAAAAAATGGCTCCCTCATTCGGTATATGCTTGGATTTTGGATTTTGGAGAGA 66
 QY 88 LeuValValSerArgTyrTrpSerIleTyrGlnAsnLeuProTrpProAspArgLeuMet 107
 Db 67 ATCGATGATACCGATGGATGGGACCAATATACCAACATTCCTTGGCCAGATGCAATTGCC 126
 QY 108 IleGlnValSerSerPheValGluGlyLysAspGluGluValArgLeuLeuArgTyrThr 127
 Db 127 ATTTGATAAAGCAGCAGATATCATCGCTCCCATGATAGGCCAGTAAATGAGCGAAT 186
 QY 128 LeuIleArgTyrAlaIleLeuGlyGlnValLeuIleLeuArgSerIleSerThrSerVal 147
 Db 187 ATCTTAATATATGTGTGCTGTGCTCCAGTATGATATTCACATGATATACCGCGTGTG 246
 QY 148 TyrLysArgPheProThrLeuHisHisLeuValLeuAlaGlyPheMetThrHisGlyGlu 167
 Db 247 AAGCGCGATTCGCCACCTACACATTAATATGAGCGCTTCTGTGTAATAATGAA 306
 QY 168 HisGlnLeuGlnLysLeuGlyLeuProHisAsnThrPhe-----TrpValPro 184
 Db 307 AAGAAATTCATGAGGCGATGATCAGGCTTCGCAAGTTATCCAAAGCATGTGATGCC 366
 QY 185 TrpValTrpPheAlaAsnLeuSerMetLysAlaTyrLeuGlyLysArgIleArgAspThr 204
 Db 367 ATCGTTTGGCAGCTGATGTGATGATGAGCGCGAGGAGAGAAATAAATCCGGATGAC 426
 QY 205 ValLeuLeuGlnSerLeuMetAsnGluValCysThrLeuArgThrGlyCysGlyGlnLeu 224
 Db 427 TACGCTGTTAAGACCATCATCGATGATGAAACAGCATTAACGGGAATGTGATTTTG 486
 QY 225 TyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValAlaThrValAlaVal 244
 Db 487 CTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
 QY 245 TyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnProAsnLysAspTyr 264
 Db 547 TACTCTCTTCT 606
 QY 265 ProGlnHisGlnMetAspLeuValValProValPheThrIleLeuGlnPheLeuPheTyr 284
 Db 607 GATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
 QY 285 MetGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGlnAspAspAsp 304
 Db 667 ATGGGTGGCTCAAGTAGCCGAGACGCTGATAAATCTTTCGAGAGGATGATGACGAC 726
 QY 305 PheGluThrAsnTrpIleIleAspArgAsnLeuGlnValSerLeuSerValAspGly 324
 Db 727 TTGACGCTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
 QY 325 MetHisGlnAsnLeuProPomMetGluArgAspMetTyrTrpAsnGluAlaAlaProGln 344
 Db 787 ATCCACACAGGAGCATCCGGAACCTGTAAGATCACTATGGGAAGAGGTCTTCCCTAAT 846
 QY 345 -----ProProTyrThrAlaAlaSerAlaArgSerArgArgHisSerPhe 359
 Db 847 GAATTCATACGCTCAGCCAAA-AATGGGTCAAAATCCTCCAGAGCTTCACAGCGCTTA 905
 QY 360 MetGlySerThrPheAsnIleSerLeuLysGlnAspLeuGluLeuTrpSerLysGlu 379
 Db 906 TATGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 965
 QY 380 GluAlaAspThrAspLysGluSerGlyTyrSerSerThrIleGlyCysPheLeuGly 399
 Db 966 GAC-----
 QY 400 LeuGlnProLysAsnTyrHisLeuProLeuLysAspLeuLysThrLysLeuLeuCysSer 419

Db 969 -----CATCTGCACACTGCTACCTACTCAGACTCC-ATC----- 1000
 Qy 420 LysAsnProLeuLeuGluGlyGlnCysLysAspAlaAsnGlnLysAsnGlnLysAspVal 439
 Db 1001 -----CAATGCAAGATGGGCCCTCCGAGTCACTTCACTGAGATCAATTTGTTTG 1054
 Qy 440 TrpLysPheLysGlyLeuAspPheLeuLysCysValProArgPheLysArgLysSer 459
 Db 1055 -----
 Qy 460 HisCysGlyProGlnAlaProSerSerHisProThrGluGlnSerAlaProSerSerSer 479
 Db 1058 CAT-----GATAGGGAACACCTCAGACACCCGTCGCATTC----- 1096
 Qy 480 AspThrGlyAspGlyProSerThrAspTyrGlnGluLysHisMetLysLysLysThr 499
 Db 1097 -----ACAATGCTTAAACGATCTAATATGACAGCGTGTGTCG--- 1135
 Qy 500 ValGluPheAsnLeuAsnIleProGluSerProThrGluHisLeuGlnAlaArgLysLeu 519
 Db 1136 -----GAAACGAGGAAACGATTCGACACCGATTTTGTG 1168
 Qy 520 AspGlnMetSerThrAsnIleGlnAlaLeuMetLysGluHisAlaGluSerTyrProTyr 539
 Db 1169 GATATGCAAAAGTACCCT-----CTGCGACATGATACCTCCGTCGTCCTTAT 1216
 Qy 540 ArgAspGluAlaGlyThrLysProVal 548
 Db 1217 AC-GATCACAGTCTTCTCTCCCCCATC 1242

RESULT 10

AAS72796
 ID AAS72796 standard; cDNA: 1526 BP.
 XX AAS72796;
 AC
 XX
 XX
 DT 13-FEB-2002 (first entry)
 XX
 XX
 DE DNA encoding novel human diagnostic protein #8600.
 XX
 XX
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dermanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-699362/73.
 XX
 DR P-PSDB; ABG08609.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID NO 8600; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 1526 BP; 404 A; 408 C; 357 G; 357 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,816-72
 Score: 742.00 Length: 1526
 Percent Similarity: 56.788 Matches: 174
 Best Local Similarity: 43.728 Conservative: 52
 Query Match: 25.208 Mismatches: 103
 DB: 23 Gaps: 69
 Gaps: 9

US-09-622-964-29 (1-551) x AAS72796 (1-1526)
 Qy 162 PheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlnProHisAsnThrPhe 181
 Db 3 TTTATGACACACAGATGAAGAAATTATTCACACACGTCAGTCTCTCATCTGAATAT 62
 Qy 182 TrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrLysGlyArgLys 201
 Db 63 TGGGTTCATCATCTGTTGGAAATCTTGCACCTAAAGCCCGAGATGAAGATGC 122
 Qy 202 ArgAspThrValLeuLeuGlnSerLeuMetLysGlnValAlaCysThrLysLeuArgLys 221
 Db 123 AGACACAGTGTTCATGATCATCATGATCATGATCATGATCATGATCATGATCATGAT 182
 Qy 222 GlyGlnLeuArgLysTrpLysTrpLysSerLysProLeuValTyrThrGlnValThr 241
 Db 183 AGCTCTTATTCGGTTTGTGACTGGTGGGATTCGCTGGTTTACACCCAGTTCGACT 242
 Qy 242 ValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnProAsn 261
 Db 243 CTGGCTGTATACCTCTCTTGGCTGCTGATGACGCCAGCTTTTGGATCCAC 302
 Qy 262 LysAspTyrProGlnHisGlnMetAspLeuValProValPheThrLysLeuGlnPhe 281
 Db 303 AAAGGCTTACAGGAGCATGCTGATCTTACATCCCATTCACCTCTACCAATTC 362
 Qy 282 LeuPheTyrMetGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyLysAsp 301
 Db 363 TTCTCTATGACAGGATGCTTAAGTACAGACAGCTTATCAACCTTTGGAGAAAT 422
 Qy 302 AspAspAspPheGlnThrAsnTrpLysLeuAspArgAsnLeuGlnValSerLeuLeuSer 321
 Db 423 GATGATGATTTTGAACATTAACGTGCATTCACAGAAATTTGCAAGCTCTCTTTACT 482
 Qy 322 ValAspGlyMetHisGlnAsnLeuProPheMetGluArgAspMetTyrTrpAsnGlnAla 341
 Db 483 GTGACAGAAATGCACATGAGCTTACCCAAAGTGAAGAGCAATTTACTGGAGCAATTC 542
 Qy 342 AlaProGlnProProTyrThrAlaAlaSerAlaArgSerArgArgHisSerPheMetGly 361
 Db 543 GCTGCTCCCAACCATACATACATGAGCTGCTGATACATACATACATACATACATACAT 602
 Qy 362 SerThrPheAsnIleSerLeuLysGlnAspLeuLysLeuTrpSerLysGlnGlnAla 381
 Db 603 TCACACAGTCCAGATGGGGCTGTCTGGGTCCGAC-----TTTCTGACGAGGAGTGG 653

```

QY 382 AsprThAspLysAspGlySerGlyTyrSerSerThrIleGlyCysPheLeuGlyLeuGln 401
DB 654 CTGTGGGATTTATGAGAGCATGCGCATCG----- 683
QY 402 ProLysAsnTyrHisLeuProLeuLysAspLysThrIleLeuLeuCysSerLysAsn 421
DB 684 -----CATTCATGATAGAGAGCAAGCCGTTCCCTGAGTCCCGCAGCAAC 731
QY 422 ProLeuLeuGlyGlyGlnCysLysAspAlaAsnGlnLysAsnGlnLysAspValTrpLys 441
DB 731 ----- 731
QY 442 PheLysGlyLeuAspPheLeuLysCysValProArgPheLysArgGlySerHis-Cy 461
DB 732 -----CCCTCCAGCCCGCAGAGAGAGAGCTACAG 761
QY 461 sGlyProGln-----AlaProSerSerHisProThrGlnGlnSerAla-ProSerSer 479
DB 762 AGCGAGACAGAGTGCACAGCTCCATGTTCTTACCC--GAGATAGCTCAGCCCGCAGG 818
QY 479 erAspThrGlyAspGlyProSerThrAspTyrGlnGlnIleCysHisMet-LysLysLys 498
DB 819 GACCTACTGATGAGCCCTCAGAAACCCCGCCAGGCGCTCAGCCAGCTGGAAGAAATCC 878
QY 499 ThrValGlnPheAsnLeuAsnIleProGlu---SerProThrGlnHisLeuGlnGln--- 516
DB 879 TGC-----TTCCAGAGAGAGAGCCCGCAGCTGATTCAGCATGGGA 920
QY 517 -----ArgArgLeuAspGlnMetSerThrAsnIleGlnAlaLeuMetLys 531
DB 921 GAGCTGTCCACCATCGAGAGACAGCCAGCAGCAACACT--TTACAGAGCGCTGACCCCA 977
QY 532 Gln--HisLacIuserTyrProTyrArgAspGlnAlaGlyThr 545
DB 978 CAGTCCAGTGTGAGAACTTCCCGCATCAAAATGCCACGCTGACC 1021

```

RESULT 11
 ABL16195
 ID ABL16195 standard; DNA; 2196 BP.
 AC ABL16195;
 XX 26-MAR-2002 (first entry)
 DT 26-MAR-2002 (first entry)
 XX Drosophila melanogaster genomic polynucleotide SBO ID NO 58.
 KW Drosophila: developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ds.
 XX Drosophila melanogaster.
 OS
 XX W0200171042-A2.
 FN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001MO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PMD, Myers EM;
 PI WPI; 2001-656860/75.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions -
 PS Claim 1: SEQ ID NO 58; 21pp + Sequence Listing: English.
 XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB53737-AB57072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX

SO Sequence 2196 BP; 546 A; 616 C; 577 G; 457 T; 0 other;

Alignment Scores:

Pred. No.:	Length:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
6.23e-69	2196	715.50	49.79%	36.42%	24.30%	177	65	158	87	12

US-09-622-964-29 (1-551) x ABL16195 (1-2196)

```

QY 78 IleSerPheValLeuGlyPheTyrValThrLeuValValSerArgTrpTrpSerGlnTyr 97
DB 1 ATGTCTTTTGTCTGGGCTCTGTATGCAATCTGTGCTGTAACGTTGGTGGAGCAATAT 60
QY 98 GlnAsnLeuProTyrProAspArgLeuMetIleGlnValSerSerPheValGlyGly--- 116
DB 61 AGCGTTTGGCTGGGCGAGATAGCTTGGCCCTGTTATAGCGCGGCTATCCCACTCA 120
QY 117 -----LysAspGlnGlyGlyArgLeuLeuArgArgThrLeuLeuArgTyrAla 132
DB 121 AATGGCGGTAAATATATATAAACGGTCCCTGATGCGCTGATATATATATATATGCGCATATG 180
QY 133 IleLeuGlyGlnValLeuIleLeuArgSerIleSerThrSerValTyrLysArgPhePro 152
DB 181 GTCTTGGCTATGTGATACCTTCAGGAAATTTCCCTTCGTTAAACGTCGTTTCC 240
QY 153 ThrLeuHisIleValLeuAlaGlyPheMetThrIleValGlnIleValGlnIleGln 172
DB 241 ACNACCCAACTGTGTGATGCGGCTGATGACAGAAATGAGATGAAATCTTGAG 300
QY 173 LysLeuGly-----LeuProHisAsnThrPheTrpValProTyrValTrpPheAlaAsn 190
DB 301 GCACGTAAATGAGAGAGCCCATGCTCCAAATATGAGATGCTCTGTTGGCTAACCAAC 360
QY 191 LeuSerMetLysAlaTyrLeuGlyGlyArgIleArgAspThrValLeuLeuGlnSerLeu 210
DB 361 ATCATTAACCGGCGCAGAAAGATGCGCTAATCGCGATCATATTGTCAAACCATTA 420
QY 211 MetAsnGlnValCysThrLeuArgThrGlnCysGlyGlnLeuTyrAlaTyrAspTrpIle 230
DB 421 CTGTAGACGCTCCGAGATATTCGTAGACGTTAGGTGCTTATTGGATGACAGACGCTC 480
QY 231 SerIleProLeuValTyrThrGlnValValThrValAlaValTyrSerPhePheAla 250
DB 481 TGTGTCTCTTGGTCTACACTCAGTAGTCACCTGCTGTACACCTATTTCATAGCT 540
QY 251 CysLeuIleGlyArgGlnPheLeuAsnProAsnLysAspTyrProGlyHisGlu---Met 269
DB 541 GCCCTTTTGGCGCGTAATGTGGCCCAAGCTTGTGACAGAGAGTACGAGATGCC 600
QY 270 AspLeuValValProValPheThrIleLeuGlnPheLeuPheTyrMetGlyTrpLeuLys 289
DB 601 GATCGTCTTCCCGCTGTTACCGATTTGCAATTTGTTTCTACGTGGGCTGAG 660
QY 290 ValAlaGlnGlnLeuIleAsnProPheGlyGlnAspAspAspPheGluThrAsnTrp 309
DB 661 GTGGCGAGAGCTTAATCAATCCCTTGGCAGAAATGAGATGATATGACCTGAACTG 720
QY 310 IleIleAspArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGlnAsnLeu 329

```

DB 721 CTAATTGACGACATCAAGCGTGCCTACATGATCGTGACGAGATGACGAGGACAC 760
QY 330 ProPromeitGuArgASpMetYrTriPasnGLuAlaAlaProGlnPro---ProYrYrTh 348
DB 781 CCCGAGCTGCTGGCGGATCTACACTGAGTGTGTGGTGGCCAGGATCTGCCCTATACG 840
QY 349 AAlaASerAlaArgSerArgArgHisSerPheMetGlySerThrPheAsnIleSerLeu 368
DB 841 GTGGCATCCGACACTACCGGAAAGACGACCGAGGCTCCCGCCAGAAATGTCACAGTTC 900
QY 369 LysLysGLuAspLeuGluLeuTrpSerLysGluGluAlaAspThrAspLysGluSer 388
DB 901 AAGAGAGAGAC----- 912
QY 389 GlyTyrSerSerThrIleGlyCysPheLeuGlyLeuGlnProLysAsnTrpHisLeuPro 408
DB 913 -----GCCATGTACGCCACATATATCCACGCGCGCGCAAGAGATG 954
QY 409 LeuLys---AspLeuLysThrLysLeuLysCysSerLysAsnProLeuGluGln 427
DB 955 CTCAGCGCAGATCTCTACCGGACATATGAACCGTGACACCGCAATGTGGAA----- 1008
QY 428 CysLysAspAlaAsnGlnLysAsnGlnLysAspValTrpLysPheLysGlyLeuAspPhe 447
DB 1009 -----AGAGAGAAAAACAACCTGGTTCGCGCAGCTC----- 1041
QY 448 LeuLysCysValProArgPheLysArgLysSerHisCysGly-ProGlnAlaProSe 467
DB 1042 -----TCCCAATGGTTCCTATAGAGACCGCAG 1068
QY 467 rSerHisProThrGlnGlnSerAlaProSerSerSerAspThrLysArgLysProSerTh 487
DB 1069 TCGAGCGCTACTCTCTCCGCGGGAATGCCATTCAACCGGACCGCTCACTCGCTTAC 1128
QY 487 rAspTrpGlnGluLysCysHisMetLysLysLysThrValGluPheAsnIleLeu 507
DB 1129 TCT-----ACTCC 1136
QY 507 cGluSer-----ProThrGlnHisLeuGlnGlnArgArgLeuAspGlnMetSerThrAs 525
DB 1137 CGAGTCCGCGCTTCCACCTGACACCTCTGACGACGACCAATTTACAG----- 1182
QY 525 nileGlnAlaLeuMetLysGlnHisAlaGlnSerTrpTrpArgAspGlnAlaGlyTh 545
DB 1183 ---CAGGCGCATCAGACGACGAAAGCGGATGCCACCA-----AG 1220
QY 545 rLysProValLeuYr 550
DB 1221 TAAATCAAGTCTCTAT 1236

RESULT 12
ABAl4558
ID ABAl4558 standard; DNA; 18537 BP.
XX
AC ABAl4558;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6889.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antistickling; antianaemic; antiathrictic; cancer;
KW antihemmatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antidiabetic; antidiabetic; antidiabetic; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN MO200159063-A2.
XX
PD 16-AUG-2001.

XX
PF 17-JAN-2001; 2001MO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198122.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224569.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226682.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249269.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 P1 Rosen CA, Barash SC, Ruben SM;
 XX WPI, 2001-541565/60.
 XX

PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 PS
 XX
 PS Disclosure; SEQ ID NO 6889; 1701pp + Sequence Listing; English.
 CC
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABAB1678-ABAB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)isogonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 2,44e-60 Length: 18537
 Score: 652.00 Matches: 313
 Percent Similarity: 19.34% Conservative: 49
 Best Local Similarity: 16.72% Mismatches: 101
 Query Match: 22.15% Indels: 1410
 DB: 22 Gaps: 20

US-09-622-964-29 (1-551) x ABA14558 (1-18537)
 QY 51 ArgMetValIleuSerSerAspGlnIleuLeuPheGluIysLeuAlaIleuTyCysasp 70
 DB 4743 AGGCTGGCCCTCAGGAAAGAACAGCTGATGTTGAAACATGACTGTATGCGAC 4802
 QY 71 SerTyrlleGlnIleuIleProlIleSerPheValIle----- 82
 DB 4803 AGCTACATCCAGCTATCCCATTTCTTCTGCTGGTGGAGTCCCTTCTGCGCTTT 4862
 QY 82 ----- 82
 DB 4863 CCGGGTCCCTGTGGCCGCCAGCCTCCAGACAGCCAGGGGAGATCAGAGAGCTGCC 4922
 QY 82 ----- 82
 DB 4923 GCAGGGGCTGGGGAGGGGGGGGGAAGCCAGCGGAGTCCGCGCTCTGTAGGG 4982
 QY 82 ----- 82
 DB 4983 AAAGTGGCGACTGCAGCCAGAGAAACTGAAGTTAGACGTTAGTAAGAGCTCTGCCGT 5042
 QY 82 ----- 82
 DB 5043 TAGCAATGAAACCCCATTTTCTAGAGGAAACCGCTGACATCATGTCCTTGAGCCCTG 5102
 QY 82 ----- 82
 DB 5103 CGCGGAGGAGGAGGGGCTGCGCGGATTTCTGGGACACAGAGGGGACCCCGGGTGACA 5162
 QY 82 ----- 82
 DB 5163 GAACCTTGGGCTCTCGCGCTTCATATGCGAGGCTGCTGCTGCTGCCAGAGCC 5222
 QY 82 ----- 82

Db 7443 CATCTCTAGTAATAAGAAATTAGTGTGTGTGTGTCAGAGCCCTGTAATCCAGCTAC 7502
QY 237 -----
Db 7503 TTGGAGAGCTGAGGAGAGAGATGCTTGAACCGGGAGGTGGAGTTGCACTGAGCTGA 7562
QY 237 -----
Db 7563 GATCATGCACTGCACCTCAGCCCTGGGCGACAGCAAGACTGTGTCTAAACAACAA 7622
QY 237 -----
Db 7623 CAACAACAACAACAACAACAAGGGGTTAACAAGCCCTTAAGTCACTAAGTG 7682
QY 237 -----
Db 7683 TGCAGTCAAGAACAGAGCCCTGTGTCTGTCTGACACTCCAGCCCTGAGACATCTG 7742
QY 238 -----
Db 7743 ATTCAAGGTTCCACCTAGCCCTTGTGACACATCCCTCTCTCTCTCTCTCTCTCC 7802
QY 238 -----
Db 238 lvalvalthrvalalavaltyrserphepheleualacysleuileglyarglnphei 258
Db 7803 AGGTGGTGCAGCTGGCGGTGTACAGCTTCTCTCTGACTTGTCTAGTTGGGGGAGATTTC 7862
QY 258 -----
Db 258 euasnprouasnllysaptyrprogluylslumetaspheleuvalprovalphei 278
QY 278 -----
Db 278 lvalleuaglupheleuphetymetelytpleu 288
QY 288 -----
Db 288 TCCTGCACTTCTTCTTCTTCTTCTTCTTCTGCTGGCTGAAGTGGGCTCTCCAGGCCCCCTG 7982
QY 288 -----
Db 7983 GCTGAGGAGATGCGCAGAGGGGTCATGCGCAGACAGCTGCTGAGAGAGATGCACTGTC 8042
QY 288 -----
Db 8043 AGGAAAGAGAGGCTCTACGGGTGTAAGAAAGCAGCCAGCGCTGTGGGCGACACCTGTATTC 8102
QY 288 -----
Db 8103 CAGCTACTCGGAGGCTGAGGAGGAGAAATCCCTTGAACCCGAGGGGAGGTGTGTGT 8162
QY 288 -----
Db 8163 GAGTTGAGATCGTCCACTGCACTCAGCTGCGGCAAAAAGAAATGAATCTATCTCAAAA 8222
QY 288 -----
Db 8223 ACAACAGACAAACAACAAGCCCTAAGTTTCAAGAACCCCTGCCCCCTTGAAGAGAG 8282
QY 288 -----
Db 8283 AGCGGACACACCTCTCTTATTAAGATGCTGTGGGCTGTCTGTCTCTCACTCAAG 8342
QY 288 -----
Db 8343 TGGCTTGTCCAGTATTCCTCCACACAGCAACCAATCTCCGAACAGATGTCTGAATCA 8402
QY 288 -----
Db 8403 CACAGTTTCTCTCCACCTCTTATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTC 8462
QY 288 -----
Db 8463 CCTTCTACCTCTCTTATTTTGTGTAATGGGGGTGAAGTCTGTGTCTGCCCCCTTC 8522
QY 288 -----
Db 8523 TGTCACTGTGACATTCCTAT 8582

QY 288 -----
Db 8583 TCCTCTAAATTCCTCCCTGCCCCCAGTATCTTGTGTTCTGACAGATCAAAACAATCA 8642
QY 288 -----
Db 8643 CACTTTATGCTTGAATTCACAGGGTGCCCGAGTGCCGTCGAAGATGTCCCTGAGC 8702
QY 288 -----
Db 8703 CCTAAGCAGACGGGTGTCACCTCTCTGCGGGCTTGTAGGCAATTTAGAGGTGCTAT 8762
QY 288 -----
Db 8763 CAGAGATCTGCCACCTAGACTGCCCTTATGTTACGCCAGCTCAGTATATCTCTG 8822
QY 288 -----
Db 8823 TTGCATGAATGAATAAATTATGCACTCCAGGTAAGATACATGAGGTGAGATAAGCA 8882
QY 288 -----
Db 8883 GTGACTCAGCCGAGTGATACACTCAGGACACAGCTGTGGGTTCAGGMAAGGACTGCTC 8942
QY 288 -----
Db 8943 AGAAGATTAGAGGGGCTGTGTCAGAAAGTGTGGGTGCCCCACAAGTGTGGGGGCTG 9002
QY 288 -----
Db 9003 AGCCCTAACTCTGCTTGAAGACAGTGTGACGAGGAGGAGGCTTCATGGGTGTGA 9062
QY 288 -----
Db 9063 AATGACAGCAGCTGAGTTTAAAGGGGAAAGCTGCTTGAAGAGTTCTGCTGAGGTT 9122
QY 289 -----
Db 9123 TACAGAGCCCTCAGCTGTCCCAAGGTGGCAGACAGCTCATCAACCCCTTGGAGAGAT 9182
QY 302 -----
Db 9183 AspaspasphegluThrAsnTPillelleasparAsnleu 315
QY 315 -----
Db 9183 GATGATGATTTTGAACCACTGATTTGTGACGAAATTTGAGGTATGGGAGAGGA 9242
QY 315 -----
Db 9243 GAGAAACATACCATGAGCTTCCCAAGTGAGACCAAGAGAGAGAGCCCACTGTCTG 9302
QY 315 -----
Db 9303 TAGGAGGCTCAGAGTGAATGATCAACCTTCCCTCTCTCTCTCTCTCTCTCTCTCT 9362
QY 315 -----
Db 9363 CACTCAGAGATTCTCACTCAATCTTGAAGCTGACAGGAGCAACCATCTCCCAATT 9422
QY 315 -----
Db 9423 CACAGGAGGAAACTGAGTTCAGAGAGAGGAGAGATTCCTTCAAGTATCAGGACA 9482
QY 316 -----
Db 9483 TACAAGTCTGCTGGGATGATCTTCTGTGGGACTTCTTCTGTCTCTCTGTCAGCAGGT 9542
QY 317 -----
Db 317 lserleuleuSerValaspelymethislnaslnleuProPromecluaRspmetty 337
Db 9543 GTCCCTGTGTGTGTGATGACATGACACAGGACCTGCTCGATGAGACCGGACATGTA 9602
QY 337 -----
Db 9603 rTTPasnGlualalalProgluProProtyrThrAlaSerAlaArgSerArgArgHt 357
Db 9603 CTGGAATTAAGCCCGAGCCACAGCCCTTACACAGTCTCTCCGCCAGTTCTGTCGAGC 9662

QY 357 sSerPheMetGlySerThrPheAsnIleSerLeuIleValGlyGluAspLeuIleuTrpSe 377
Db 9663 CTCTTTATGGGCTCCACACTTCACATCAGTGTGGCCAGAGCCAGGGGGCTGGTGGGA 9722
QY 377 rLyGluGluAlaAspThrAspIlySlyGluSerGlyTyr-----SerSe 392
Db 9723 AGCCCCCTCAGT-----GCAGGGGCTGGCTAGGAACCTAGAAATAG 9764
QY 392 rThrIleGlyCys-----PheLe 398
Db 9765 CACTAGTTAATGATACAGGTGCTTCAGTAAGTCAGAGCACTAGTATGCTCTTAT 9824
QY 398 uGlyLeuGlnProIlyAsnTyrHisLeuPro----- 408
Db 9825 AATCATT-----AACTATTTTCTCCCAATAATTCGGTTTGTATCCCAAGT 9875
QY 409 -LeuIlyAspLeuIlyThrIlyLeuLeuGlySerIlyAsnPro----- 422
Db 9876 TTTACGATTAATTAAGTACAGTTCAGAGAGAGTAAAGTTGTCAGAGCCAGATAGTACC 9935
QY 423 -----LeuLeuGluGlyGln-----CysIlyAspAlaAsnGlnIlyS 434
Db 9936 AATGTGTCATTGTCCTACTCGAGAGCAGCCTATGATCAGTACAGTACAGTACAGT 9995
QY 435 -----AsnGlnIlyAspValTrpIlySphIlyGlyLeuAspPheLeuIlySc 450
Db 9996 ACCTGCTCTTGTATCCAGAACATATGTTTCTT-TCTTTTGGAGACATCTCGCTC 10054
QY 450 yValProIlySphIlyAspArgIly-----SerHisCysGlyProGlnAlaProS 467
Db 10055 TGTGCCCGAGTGGTGGCGGCGATCTTGCGTACCTGCAACCTCCGCTCTG 10114
QY 467 eSerHisProThrGluGlnIleSerAla---ProSerSerSerAspThrGlyAspIlyProS 486
Db 10115 GTTCAAGATGATTCCTCGCTTACACCTCCCGAGTACGCTGGGATTAACAGT----- 10164
QY 486 eThrAspIlyrGlnGlnIleCysHisMetIlySlySlyThrValGluPheAsnLeuAsnI 506
Db 10165 -----GCCCAACCAACCAACAGCGCAATT---TTTGACTT---T 10198
QY 506 leProGluSerProThrGluHisLeuGlnIleArgIlyLeuAspGlnMetSerThrAsnI 526
Db 10199 TAGTAGAGATGAGTTCACACATGTGGCCAGGCTGTCTCCAACTCTGACCACTAATC 10258
QY 526 leGlnAlaLeuMetIlyGluHisAla 534
Db 10259 TGCCCGCTTGGCTCCCAAAATGCT 10284

RESULT 13
ID ABA14559 standard; DNA; 16650 BP.
AC ABA14559;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6890.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antitubercial; vulterary;
KW antiparkinsonian; antischizking; antianemic; antitubercial; cancer;
KW antitubercial; hepatotoxic; cerebroprotective; antitubercial;
KW antitubercial; antidiabetic; antitubercial; antitubercial;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX
PN MO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-USO1334.

XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 30-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228927.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232198.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.

02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249220.
 PR 17-NOV-2000; 2000US-0249221.
 PR 17-NOV-2000; 2000US-0249222.
 PR 17-NOV-2000; 2000US-0249223.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249226.
 PR 17-NOV-2000; 2000US-0249227.
 PR 17-NOV-2000; 2000US-0249228.
 PR 17-NOV-2000; 2000US-0249229.
 PR 17-NOV-2000; 2000US-0249230.
 PR 17-NOV-2000; 2000US-0249231.
 PR 17-NOV-2000; 2000US-0249232.
 PR 17-NOV-2000; 2000US-0249233.
 PR 17-NOV-2000; 2000US-0249234.
 PR 17-NOV-2000; 2000US-0249235.
 PR 17-NOV-2000; 2000US-0249236.
 PR 17-NOV-2000; 2000US-0249237.
 PR 17-NOV-2000; 2000US-0249238.
 PR 17-NOV-2000; 2000US-0249239.
 PR 17-NOV-2000; 2000US-0249240.
 PR 17-NOV-2000; 2000US-0249241.
 PR 17-NOV-2000; 2000US-0249242.
 PR 17-NOV-2000; 2000US-0249243.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249246.
 PR 17-NOV-2000; 2000US-0249247.
 PR 17-NOV-2000; 2000US-0249248.
 PR 17-NOV-2000; 2000US-0249249.
 PR 17-NOV-2000; 2000US-0249250.
 PR 17-NOV-2000; 2000US-0249251.
 PR 17-NOV-2000; 2000US-0249252.
 PR 17-NOV-2000; 2000US-0249253.
 PR 17-NOV-2000; 2000US-0249254.
 PR 17-NOV-2000; 2000US-0249255.
 PR 17-NOV-2000; 2000US-0249256.
 PR 17-NOV-2000; 2000US-0249257.
 PR 17-NOV-2000; 2000US-0249258.
 PR 17-NOV-2000; 2000US-0249259.
 PR 17-NOV-2000; 2000US-0249260.
 PR 17-NOV-2000; 2000US-0249261.
 PR 17-NOV-2000; 2000US-0249262.
 PR 17-NOV-2000; 2000US-0249263.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249266.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249268.
 PR 17-NOV-2000; 2000US-0249269.
 PR 17-NOV-2000; 2000US-0249270.
 PR 17-NOV-2000; 2000US-0249271.
 PR 17-NOV-2000; 2000US-0249272.
 PR 17-NOV-2000; 2000US-0249273.
 PR 17-NOV-2000; 2000US-0249274.
 PR 17-NOV-2000; 2000US-0249275.
 PR 17-NOV-2000; 2000US-0249276.
 PR 17-NOV-2000; 2000US-0249277.
 PR 17-NOV-2000; 2000US-0249278.
 PR 17-NOV-2000; 2000US-0249279.
 PR 17-NOV-2000; 2000US-0249280.
 PR 17-NOV-2000; 2000US-0249281.
 PR 17-NOV-2000; 2000US-0249282.
 PR 17-NOV-2000; 2000US-0249283.
 PR 17-NOV-2000; 2000US-0249284.
 PR 17-NOV-2000; 2000US-0249285.
 PR 17-NOV-2000; 2000US-0249286.
 PR 17-NOV-2000; 2000US-0249287.
 PR 17-NOV-2000; 2000US-0249288.
 PR 17-NOV-2000; 2000US-0249289.
 PR 17-NOV-2000; 2000US-0249290.
 PR 17-NOV-2000; 2000US-0249291.
 PR 17-NOV-2000; 2000US-0249292.
 PR 17-NOV-2000; 2000US-0249293.
 PR 17-NOV-2000; 2000US-0249294.
 PR 17-NOV-2000; 2000US-0249295.
 PR 17-NOV-2000; 2000US-0249296.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249298.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251879.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259676.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-541565/60.
 DR
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system

PT cancers and metastases -
 XX Disclosure: SEQ ID NO 6890; 1701pp + Sequence listing; English.
 PS
 CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
 CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (anti)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 other;
 Alignment Scores:
 Pred. No.: 4,43e-60 Length: 16650
 Score: 649.00 Matches: 333
 Percent Similarity: 19.36 Conservative: 49
 Best Local Similarity: 16.74 Mismatches: 101
 Query Match: 22.04 Indels: 1408
 DB: 22 Gaps: 20
 US-09-622-964-29 (1-551) x ABA14559 (1-16650)
 QY 51 ArgmetValIleuSerSerAspInIleuLeuPheGluYlYleuAlaLeuTyCysAsp 70
 DB 4743 AGGCTGGGCGCTACGAGAAACACAGCTGATGTTGAGAAACAGCTGTATTTGGCAG 4802
 QY 71 SerTyrlIleGlnIleuIleProIleSerPheValIleu
 DB 4803 AGTACATCCAGCTCATCCCATTTCTTGTGTGGTGAGTTCCCTTTCGGCTGT 4862
 QY 82
 DB 4863 CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGCCAGGGAGAGATCAGAGAGCTGCG 4922
 QY 82
 DB 4923 GCAGGGGCTGGGAGGGGGCGGGGGAAGCCAGCGGAGGTGGCGCTCTGTAGGG 4982
 QY 82
 DB 4983 AAGGTCGGAGCTGCAGCCAGAGAAACTGAGTTAGAGTTAGTAAGAGCTCTGCGCT 5042
 QY 82
 DB 5043 TAGCAATGAACCCCATTTTCTGAGGAGGAGCGCTGACATCATGTCTCTGGAGCCCTG 5102
 QY 82
 DB 5103 CCGGGAGGGGAGGGGGTCTGCGGATTTCTGAGACAGAGGGGAGACCCCGGGGTGACA 5162
 QY 82
 DB 5163 GAACCTTGGGGCTCTCGGGCTCCATCGAGGCTCTGCTCTGCTCCGAGCGC 5222
 QY 82
 DB 5223 CTTCCAGAGGAGGCTGGGGCTAGAGCCGCTCGCAGAGAAAGCTGGAGAGAGCGAGCAT 5282
 QY 82

02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239933.
 PR 20-OCT-2000; 2000US-0240966.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0242221.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249287.
 PR 17-NOV-2000; 2000US-0249289.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251997.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX

PS Disclosure; SEQ ID NO 6888; 1701bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABAI1004-ABA21534) and proteins
 CC (ABAI14678-ABAI18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIFO at ftp.wifo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 other;
 Alignment Scores:
 Pred. No.: 5 28e-60 Length: 18530
 Score: 649.00 Matches: 313
 Percent Similarity: 19.36% Conservative: 49
 Best Local Similarity: 16.74% Mismatches: 101
 Query Match: 22.04% Indels: 1408
 DB: 22 Gaps: 20
 US-09-622-964-29 (1-551) x ABAI14557 (1-18530)
 QY 51 ArgMetValLeuSerSerAspGlnGlnLeuLeuPheGlnLysLeuAlaLeuTyrcysAsp 70
 DB 4743 AGGCTGGCCCTCAGCGAGGAGAACACAGCTGATGTTGAGAACGACTCTGTTGGAC 4802
 QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
 DB 4803 AGCTACATCCAGCTCATCCCCATTTTCTCTGCTGGTGATGCCCTTCCTGCTGTT 4862
 QY 82 ----- 82
 DB 4863 CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGGCCAGGGAGATCAGAGAGCTGGC 4922
 QY 82 ----- 82
 DB 4923 GCAAGGGCTGGGGAGGGGGGGGAGGAGCGACGCGAGCTGGCGCTCTGTAGGG 4982
 QY 82 ----- 82
 DB 4983 AAGGTGCGGACTGACGACAGAGAAACTGAAGTTAGACTTAGTAAGACTCTGCGGT 5042
 QY 82 ----- 82
 DB 5043 TAGCAATGAAGAACCCCATTTTCTGAGGAGAGCGCTGACATGATGTCCTGAGCCCTG 5102
 QY 82 ----- 82
 DB 5103 CGCGGAGGAGGAGGGGCTGTGGCGGATTTCTGGGACAGCAGGGGAGACCCCGGGTGACA 5162
 QY 82 ----- 82
 DB 5163 GAACCCCTGGGGCTCTCGCGCTCCATGACGAGGCTTGCCTCTGCTCCGAGAGCG 5222
 QY 82 ----- 82
 DB 5223 CTTCAGAGAGGCTGGGGGCTAGGCCCGCTCCGACGAAAGCTGGAGAGCCGAGGCAT 5282
 QY 82 ----- 82
 DB 5283 CGCGGGCGCTGGGGCTGTGGCGGACGCTGGGCCCTCGCCCTCGCCCGCCGCG 5342

Qy	83	-----GlyPheTyrValThrLeuValValSerArgTrrPrrSerGlnYrGlu	98
Db	5343	CCCTCCGCCAGGCTTCTACGTGACGCTGTGTGACCCGGGTGGAAACCAAGTACGAG	5402
Qy	99	AsnLeuProTrrPrrAspArgLeuMetIleGlnValSerSerPheValJluIyAsp	118
Db	5403	AACCTGCCGTGGCCCGCCCATGACGCTGTGTGACCCGGGTGGAAACCAAGTACGAG	5462
Qy	119	GluGluValTrrLeuLeuArgTrrLeuIleArgTyrAlaIleLeuGluValLeu	138
Db	5463	GACCAAGCCGCGCTGTGTGACGCTGTGTGACCCGGGTGGAAACCAAGTACGAG	5522
Qy	139	IleLeuArgSerIleSerThrSerValTrrLysArgPrrProThrLeuHisIleVal	158
Db	5523	ATCCTCGCAGCGGTGACGACCCGACGTCTCAAGCCCTTCCAGGCCACCACTGTG	5582
Qy	158	-----	158
Db	5583	CAAGCAGTGGGCGGACCGGAGCAACGGGAGGACCGGCGGACGAGGCCGAGAT	5642
Qy	158	-----	158
Db	5643	GGGCGCGGAGGAATGAAAGATGGGTGAGCCAAAGTCCCGGACTCGGGGATTTGGT	5702
Qy	158	-----	158
Db	5703	GGAGCCAGAGTGGGGTGTGTCAAGATTGGGGGTCCAAATGGGGCGGACAGATCGG	5762
Qy	158	-----	158
Db	5763	TGTCTGAAGGTGGGGCGAGGCAAGGACCCACCTCCGAGATGAGATCTGAGGACGG	5822
Qy	158	-----	158
Db	5823	CTAAGACCTTGAGGATATGAAAGAAGGCTGAGGCTTGGAACTGTGAGATCT	5882
Qy	158	-----	158
Db	5883	AGGCTACTTCCCTGTGCCCTTGCCCTTGTATCTCGGTTTCCACTTGGAGTATG	5942
Qy	158	-----	158
Db	5943	GGACATTGGTCTGTGACACCCCTCAGCCTGAGCTGTGTCTGTATTAAGACAG	6002
Qy	158	-----	158
Db	6003	ACCAAGGCTAGGCGTGTGTCTGTGCTGTATCCAGTGTCTTAGAGGCAAGGTGG	6062
Qy	158	-----	158
Db	6063	GAAGATCGCTGAGCCAGCTGTTTGAGACGCCCTTGACAAATAGCAGACCCCATC	6122
Qy	158	-----	158
Db	6123	TCTACAAAACATTAAAAATTAGCAGGGCATGTGTGGTGTGTGTGTGTGTGTGA	6182
Qy	158	-----	158
Db	6183	GTATCGGAGGCTGAGGACAGAGATCACTTGAAGCCGACGATTCAGGCTGACGTGG	6242
Qy	158	-----	158
Db	6243	TAAAGTGCACCGCTGCATCTCAACCTCGGTGACAGAGCCACCTTCTCTGGAATA	6302
Qy	158	-----	158
Db	6303	AATAAATACCTGCCCATGCTTCAGCCAGAACACCACTAGTAGTGTCTAGAAATTT	6362
Qy	158	-----	158
Db	6363	TTTTGTGTGAAGAAGAGATGCAAGAGAGTGTCTAGTCTCTATAGGTGACGAG	6422
Qy	159	-----Leu-AlaG1 161	161

Db	6423	TGCGGCCATCCCTTCTGCAGTCTTCCACCCACCGCCCTTCTACTCTCAGCTGACG	6482
Qy	161	YrPheMetThrHisGlyGluHisLysGlnLeuGlnLysLeuGlyLeuProHisAspThrPh	181
Db	6483	CTTATGACTCCCGCAGAACACACAGCAGTGGAGAAACTGACCTTACCAACATGTT	6542
Qy	181	eTrrValProTrrValTrrPrrPheAlaAsnLeuSerMetLysAlaTrrLeuGlyArgI1	201
Db	6543	CTGGTGGCCCTGGGT	6602
Qy	201	eArgAspThrValIleLeuGlnSerLeuMetAsn	212
Db	6603	CCGGACCTATCTGTCTCAGAGCCCTCTGAACTGAGGCCACTGTACAGACAGGCTG	6662
Qy	212	-----	212
Db	6663	CCGACAGTGGGAAGGT	6722
Qy	212	-----	212
Db	6723	GGGCGCTGAGGCTTCCAGACCCGAGGTGGGTTGCAAGATCTTTCCACACGAA	6782
Qy	213	-----GluValCys-----	215
Db	6783	TCCACAGCCGAGGT	6842
Qy	215	-----	215
Db	6843	CTGTGTTCCCTTTGATAGATGAGAAAGCTGAGACACAAAGGTTTATGACTTCCA	6902
Qy	215	-----	215
Db	6903	TGGCACACAGCCAGGAATGACCATAGTACCAGGCCCTGTACTGTGAGAGAGTGG	6962
Qy	215	-----	215
Db	6963	GGGCGAGCCAGGTTGGGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7022
Qy	216	-----ThrLeuArgThrGlnCysGlyGlnLeuTrrAlaTrrAspTrrIleSerI1	232
Db	7023	AGGAGATGAACACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7082
Qy	232	IleProLeuValTrrThr-----	237
Db	7083	TCCCACTGTGTATACAGAGTGAAGACTAGGCTGTGTGTGTGTGTGTGTGTGTGT	7142
Qy	237	-----	237
Db	7143	AGGCTAGAAGACCAAGAGACAGCTGGGGTGGGAAGGCTCACCTAGAGCTAAGTGC	7202
Qy	237	-----	237
Db	7203	TCCCTGGGAGTGGGT	7262
Qy	237	-----	237
Db	7263	AAGATCAGGCTCTGT	7322
Qy	237	-----	237
Db	7323	AAGAGAGTGGCCGGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7382
Qy	237	-----	237
Db	7383	GGCAGTGTATCACCTGAGGTGAGGTTCGAGACCAAGCTTGGCAACATGTGTAAACC	7442
Qy	237	-----	237
Db	7443	CATCTCTAGAAATACAGAAATTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	7502
Qy	237	-----	237


```

Db 7503 TTGGAGAGCTGAGGAGAGAAATCGCTTGACCCGGAGGTGAGGTGACGTGACGTGA 7562
QY 237 -----
Db 7563 GATCATGCCACTGCACTCCAGCCTGGGCGACACAGCAAGACTCTGTCTCAAAACAACAAA 7622
QY 237 -----
Db 7623 CAACAACAACAACAAGGGGTTACAGAGCCCTTAAGTCAATAGTGTGCAAGTC 7682
QY 237 -----
Db 7683 AGAACAGGCTTGGTCTCTCTCTCAGACTCCAGCCCTGAGCATCTGATTTCAGG 7742
QY 238 -----
Db 7743 GTTCCACCTAGCCCTTGTCTACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7802
QY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgIlePheLeuAspPro 260
Db 7803 ACTGGGGGGGTGACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7862
QY 261 AsnLysAspTyrProGlyHisGluMetAspLeuValProValPheThrIleLeuGln 280
Db 7863 GCCAAGGCTACCTCGGCACTGACCTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7922
QY 281 PheLeuPheTyrMetGlyTyrPleu-----
Db 7923 TTCTCTCTATGTTGGCTGGCTGAAGTGGCCCTCTCCAGGGCCCTGCTGGGCTGGAAG 7982
QY 288 -----
Db 7983 CATGCCAGAGGGGTCTATGCCACAGCACTGCTGAGACGAGATGACGTCTCAGAAAG 8042
QY 288 -----
Db 8043 AAGGTCTACGGGTAGAAAAGCAGCAGCGCGTGTGGCGCACACCTTAATCCAGCTACT 8102
QY 288 -----
Db 8103 CGGAGAGCTGAGGAGAGAAATCGTTGAACCCGGAGGAGGTGTGTGATGTGAG 8162
QY 288 -----
Db 8163 ATCGTCACACTGCACTCCAGCCTGGGCAAAAGATGAATCTATCTCAAAAACAACAAC 8222
QY 288 -----
Db 8223 AACACAAAACAAGCCCTTAAGTTTCAGAAAGCCCTGCGCTTTAGAAAGCAGAGGGGAC 8282
QY 288 -----
Db 8283 CACCTCTCTTATTCAGATGCGCTGTGGCTGTCTTCTTCTCACTCAAGTGGCTGTG 8342
QY 288 -----
Db 8343 CCAGGTATTCCTCCACACAGCAATATCTCGAAGAGATGTTCTGATACACAGATT 8402
QY 288 -----
Db 8403 TCTCTCACCTTTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8462
QY 288 -----
Db 8463 CTTCTCTTATTTTGTATAGGGGTGTAAAGTCTGTCTGTCTGCGCCCTCTCTCTCTCTCT 8522
QY 288 -----
Db 8523 TGACACACACACACACACACACACACACACACACACACATTCCTATTCCTCT 8582
QY 288 -----
Db 8583 AAATTCCTCTGCCCCCAGTATCTTGTCTTCTGCAAGATCAAAAACAATCACACTTT 8642
QY 288 -----

QY 288 -----
Db 8643 TATGCTGAATTTCTCCAGGTGCCCCAGTGGCTGCAAGATGTCCCTTGACCCCTAAG 8702
QY 288 -----
Db 8703 GCAGAGCGGTACCTCTTCGGGGCTTTGTTAGGCAATTTAGAGTTTCTATCCAGGA 8762
QY 288 -----
Db 8763 ATCTGCCACTAGACTGCCCTTTAGTTACGCCAGCTTCAATATATCTCTGTTCAT 8822
QY 288 -----
Db 8823 GAATGAATAAATTTATGCAATCTCAGGTAAAGTACATGAGTGAATAAAGCAGTACT 8882
QY 288 -----
Db 8883 CAGCCGAGTATACACTCAGGAGACAGCTGTGGTGTTCAGGAAAGACTGGCTCAGAGA 8942
QY 288 -----
Db 8943 GTTAGAGGCGCTGTGTCCAGAAAGTGTGTGGTCCCAAGTGTGGGGGCTGAGGCCCT 9002
QY 288 -----
Db 9003 AAACCTGCTTTGAAACAGCTGTCAGGACAGAAAGGCTTCATGGGGTGTGAAATAGC 9062
QY 288 -----
Db 9063 AGCAGCTGAGGTTTAAAGGGGAAAGTGGCTTTGAGGAGTTCGTCTGAGGTTTACAGA 9122
QY 289 -----
Db 9123 GCTTCACCTGTCCCAAGGTGGCAGACACCTATCAACCTTTTGGAGAGATGATGAT 9182
QY 304 AspheIuThrasnTrpIleAspArgAsnLeu-----
Db 9183 GATTTGAGCAACATGGATTGTGCACAGAAATTCAGAGTATGGGAGAGAGAGAAA 9242
QY 315 -----
Db 9243 CCATACATGAGCTTCCCAAAAGTGACCCAAAGAGAGACCACACTGTCTGTAGGGA 9302
QY 315 -----
Db 9303 GGCTCAGAGTAAATGATCAACCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9362
QY 315 -----
Db 9363 CAGGATTCACCTCAATCTTTAGAGCTGACAGGACGCCATCTCCCATTTACAGG 9422
QY 315 -----
Db 9423 CAGGAAACTGAGTCCAGAGAGAGAGAGATTCCTCAAGTCAAGCCATACAG 9482
QY 316 -----
Db 9483 GTTCCTGCTGGGATATCTTCTGTGGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9542
QY 319 uLeuSerValAspGlyMetHisGlnAsnLeuProPheMetIuArgAspMetTyrTrpAs 339
Db 9543 GTTGGCTGTGATGATGACACAGGAGCTCTCGATGAGACCGGACATGATCTGAA 9602
QY 339 nGluAlaIleProGlnProPheTyrHisAlaIleSerAlaArgSerArgHisSerPh 359
Db 9603 TAAAGCCGAGCAGAGCCCTCTACACAGCTCTTCCGCGCAGTTCGGTGGAGCTCTCT 9662
QY 359 eMetGlySerThrPheAsnIleSerLeuLysGluAspLeuGluLeuTrpSerIysG 379
Db 9663 TATGGGCTCCACCTTCAACATCAGGTGTGGCCAGAGCAGGAGGCGGTGGTGGAGAGCCCC 9722
QY 379 -----
```

```

OY 379 uGlnAlaAspThrAspLysLysGluSerGlyTyr-----SerSerThrI 394
Db 9723 TCCTAGT-----GCAGGGGTCTGCCCTAGGAACCTAGTAATACACAG 9764
OY 394 eGlyGys-----PheLeuGlyLe 400
Db 9765 TTAATGATACAGAGTTCCTCAGTAAGTCAGACGACGCTACTACTTATTAACAT 9824
OY 400 uGlnProLysAsnThrHisLeuPro-----LeuLy 410
Db 9825 T-----AACCTATTTTCTCCCAATATTCGTTTGTATCCCAAGTTTCAG 9875
OY 410 sAspLeuLysThrLysLeuLeuCysSerLys-AsnPro-----422
Db 9876 ATAAATTAAAGTACAGGTTCCAGAGAGACTAGTCTCCAGGCCACATAGCTACCAATAG 9935
OY 423 -----LeuLeuGluGlyGln-----CysLysAspAlaAsnGlnLys-----434
Db 9936 TGCATTGCTACTCTGGAAGGACACCCATATGATCAGTATGACATGAGAAAGTTAGACCTGG 9995
OY 435 -----AsnGlnLysAspValATrPlyPheLysGlyLeuAspPheLeuLysCysValP 452
Db 9996 CTCTTGTCATCCAGAACTATAGTTTCTT-TTCTTTTGAGACAGTATCTCGCTCTGTCGC 10054
OY 452 roArpPheLysArgArgGly-----SerHisCysGlyProGlnAlaProSerSerH 469
Db 10055 CCAAGTTGGAGCCAGCTGCGTGTATCTTGCTCAGCTACGCAACCTCCGCCCTCCGGGTTCAA 10114
OY 469 IsPProThrGluGlnSerAla-----ProSerSerSerAspPheLysArgLysProSerThrA 488
Db 10115 GTGATTCTCTCTGCTTCCAGCTCCGCCAGTTCGATTCAGGT-----10158
OY 488 sPTrGlnGluLeuLeuGlyHisMetLysLysLysThrValGluPheAsnLeuAsnLeuProG 508
Db 10159 -----GCCACAAACACAACTGCTAATT-TTGTACTT---TTAGAG 10198
OY 508 IsPProThrGluHisLeuGlnGlnArgArgLeuAspGlnMetSerThrAsnLeuGlnA 528
Db 10199 AGATGAGGTTTCACCATGTGTGGCCAGCGCTGCTCTCAACTCCTGACAGTAATCTGCCCG 10258
OY 528 IsLeuMetLysGlnHisAla 534
Db 10259 CTTTGCCCTCCCAAAATGCT 10278

RESULT 15
AAF76848
ID AAF76848 standard; cDNA; 1350 BP.
AC XX
AAAF76848;
XX XX
14-MAY-2001 (first entry)
DE XX
Human secreted protein cDNA #6.
XX XX
Human; secreted protein; immunomodulatory; antisclerotic;
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW immunostimulant; cytostatic; cardiant; vasculant; anti-angiogenic;
KW antitumorological; neuroprotectant; neurotropic; anticonvulsant; vaccine;
KW antitumorals; antiparkinsonian; antimicrobial; vulnerrary; gene therapy;
KW immune disorder; hyperproliferative; cardiovascular; angiogenic;
KW neurological; infection; ss.
XX XX
Homo sapiens.
OS XX
PN W0200112776-A2.
PD 22-FEB-2001.
XX XX
15-AUG-2000; 2000MO-US22350.
PF XX
16-AUG-1999; 99US-0148759.
PR XX
PA (HUMA-) HUMAN GENOME SCI INC.

```

XX Sht Y, Young PE, Edner R, Soppet DR, Ruben SM;
 PT WPI: 2001-244245/25.
 DR P-PSDB; AAB70067.
 XX
 PT Nucleic acids encoding 18 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 XX disease and diabetic retinopathy -
 XX
 PS Claim 1: Page 352; 380pp; English.

The present sequence is one of 18 nucleic acid molecules encoding novel
 CC human secreted proteins. The nucleic acids and proteins may be used in
 CC the prevention, diagnosis and treatment of diseases including immune
 CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and
 CC human immunodeficiency virus (HIV) infections), hyperproliferative
 CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. sclerotic syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration
 CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples. The polypeptides may also be
 CC used as antigens in the production of antibodies and in assays to
 CC identify modulators of protein expression and in assays to
 XX

Sequence 1350 BP; 274 A; 417 C; 363 G; 277 T; 19 other;

Alignment Scores:
 Pred. No.: 5,05e-61 Length: 1350
 Score: 641.50 Matches: 117
 Percent Similarity: 53.854 Conservative: 42
 Best Local Similarity: 41.888 Mismatches: 86
 Query Match: 21.798 Indels: 76
 DB: 22 Gaps: 10

US-09-622-964-29 (1-551) x AAF76848 (1-1350)

```

OY 179 AsnThrPheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrLeuGly 198
Db 59 AACAGTACTGGGCTCCCTGCGCTGTGTTCAACCACTGGCGCCAGCCGAGGGAC 118
OY 199 GlyArgIleArgAspThrValLeuLeuGlnSerLeuMetAsnGluValCysThrLeuArg 218
Db 119 GGGCAATATACGAGCATATCGCTCTGCTACTTTTGGAAAGAGCTAACCACTACCGA 178
OY 219 ThrGlnGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGln 238
Db 179 GCCAAGTCCAGCAGTATTCCTACTATGATGATGATGATGATGATGATGATGATGATGATG 238
OY 239 ValValThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeu 258
Db 239 GTGGTGCACATATGCGGCTCTACTCTTTTGGCTCTCCCTGCTGCTGCTGCTGCTGCTGCT 258
OY 259 AsnProAsnLysAsp-----TyrProGlnHisGlu-----268
Db 299 GAGCCAGAGGCGAGGGCTGCCAAACCTCAGAACCTTGTGAAGCCAGGACGAGCCAGCC 358
OY 269 -----MetAspLeuValValProValPheThrIleLeuGlnPheLeuPhe 283
Db 359 CCAAGCCCTGGAGACCCGCGACATGTCAGTGCCTCTCCACCATCTGCTGACGTTCTTCTTC 418
OY 284 TyrMetGlyTrpLeuLysValAlaGluGlnLeuLeuAsnProPheGlyGlnAspAsp 303
Db 419 TATGCTGGCTGCTCAAGGTGCTGTAACAGATCATCAACCACTTGGTGAATGATGATGAC 478
OY 304 AspPheGluThrAsnTrpIleIleAspArgAsnLeuGlnValSerLeuLeuSerValAsp 323
Db 479 GACTTGGACAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 538

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model
Run on: July 27, 2003, 11:07:41 ; Search time 4217.55 Seconds
(without alignments)
3802.121 Million cell updates/sec

Title: US-09-622-964-29
Perfect score: 2944
Sequence: 1 MTITVTKVANKRGSRSSL.....EHAESYPRDEACTKPLVYE 551

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n.model -DEV-xlh
-Q/cgcn2.1/USPNO/US09622964/runat_22072003.101118_26013/app_query.fasta_1.2069
-DB-GenEmbl1 -QFAST-fastap -SUFFIX-rge -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-Blosum62 -TRANS-human40.cdi -LIST-45
-OUTFILE-pio -NORM-ext -HEAPSIZE-500 -MINLEN-0 -ALIGN-15 -MODE-LOCAL
-USER-US09622964.ecgn.1.1.7526.ecgnat.22072003.101118_26013 -NCPU-6
-NO_WMAP -LARGESUBSTRY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : GenEmbl1.*
1: gb_da.*
2: gb_htg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1910	64.9 2171 9	AF073501	AF073501 Homo sapi
2	1910	64.9 2210 9	AF057169	AF057169 Homo sapi
3	1841.5	62.6 2420 9	AF057170	AF057170 Homo sapi
4	1285	43.6 2500 9	AK096459	AK096459 Homo sapi
5	1261.5	42.8 1506 10	AF440758	AF440758 Homo sapi
6	1228.5	41.7 1956 10	BC031186	BC031186 Mus muscu
7	1227.5	41.7 1957 10	BC019528	BC019528 Mus muscu
8	1227.5	41.7 1957 10	BC031186	BC031186 Mus muscu
9	1168.5	39.7 2045 9	AF440756	AF440756 Homo sapi
10	1135	39.2 1289 4	AF440757	AF440757 Homo sapi
11	1147.5	39.0 1861 9	AK000139	AK000139 Homo sapi
12	961	32.6 2874 3	AF218817	AF218817 Homo sapi
13	961	32.6 2874 3	AF061546	AF061546 Drosophila
14	829	28.2 32022 2	AC019521	AC019521 Drosophila
15	829	28.2 170675 3	AC008139	AC008139 Drosophila
16	829	28.2 192055 3	AC009183	AC009183 Drosophila
17	829	28.2 221888 3	AC009183	AC009183 Drosophila
18	826	28.1 125804 2	AC019853	AC019853 Drosophila
19	824.5	28.0 163514 3	AC091227	AC091227 Drosophila
20	824.5	28.0 274351 3	AE003521	AE003521 Drosophila
21	774.5	26.3 125804 2	AC019853	AC019853 Drosophila
22	774.5	26.3 163514 3	AC019853	AC019853 Drosophila
23	774.5	26.3 274351 3	AE003521	AE003521 Drosophila
24	699	23.7 90487 9	AL592166	AL592166 Human DNA
25	679.5	23.1 147750 9	AC018761	AC018761 Homo sapi
26	674.5	22.9 22973 3	AF025458	AF025458 Caenorhabd
27	652	22.1 112309 9	AF139813	AF139813 Human Chr
28	649	22.0 142092 9	AF139813	AF139813 Human Chr
29	649	22.0 196080 3	CHRG39K23	CHRG39K23 Homo sapi
30	644.5	21.9 39590 3	AC004228	AC004228 Homo sapi
31	642	21.8 163815 2	CHRG39K23	CHRG39K23 Homo sapi
32	642	21.8 163815 2	AC087451	AC087451 Homo sapi
33	633	21.5 133683 2	AF003733	AF003733 Homo sapi
34	612.5	20.8 160169 2	AC084857	AC084857 Homo sapi
35	600.5	20.4 22973 3	AC051664	AC051664 Homo sapi
36	591.5	20.1 330724 2	AF025458	AF025458 Caenorhabd
37	590.5	20.1 39333 3	CEB0564	CEB0564 Caenorhabd
38	590	20.0 181453 2	AC130166	AC130166 Homo sapi
39	573.5	19.5 215043 2	AC073812	AC073812 Homo sapi
40	549	18.6 37977 3	L16621	L16621 Mus muscu
41	546	18.5 17402 3	CEC29F4	CEC29F4 Homo sapi
42	525.5	17.8 27394 3	CEB32B6	CEB32B6 Homo sapi
43	517	17.6 48095 3	CET20G5	CET20G5 Homo sapi
44	511.5	17.4 105126 2	AC014828	AC014828 Drosophila
45	511.5	17.4 171209 3	AC091202	AC091202 Drosophila

RESULT 1

ALIGNMENTS

```

AF073501
LOCUS       AF073501                2171 bp    mRNA    linear    PRI 02-SEP-1998
DEFINITION complete cds.
ACCESSION   AF073501
VERSION     AF073501.1  GI:3511241
KEYWORDS
SOURCE      Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 2171)
AUTHORS     Stohr, H., Margardt, A., Rivera, A., Cooper, P. R., Nowak, N. J.,
            Shows, T. B., Gerhardt, D. S. and Weber, B. H.
            A gene map of the Best's vitelliform macular dystrophy region in
            chromosome 11q12-q13.1
            Genome Res. 8 (11): 48-56 (1998)
JOURNAL     98112782
MEDLINE     9445487
PUBMED
REFERENCE   2 (bases 1 to 2171)
AUTHORS     Margardt, A., Stohr, H., Passmore, L., Kraemer, F., Rivera, A. and
            Weber, B. H.
            Direct Submission
            Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am
            Hubland, Wuerzburg 97074, Germany
FEATURES
     source          1..2171
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11q12-q13.1"
     gene            1..2171
                     /tissue_type="retinal pigment epithelium"
     CDS             89..1846
                     /gene="VMD2"
                     /note="Best disease"
                     /product="vitelliform macular dystrophy protein"
                     /protein_id="AAC33766.1"
                     /db_xref="GI:3511242"
                     /translation="MTTTSVAVNARLGSFRLCLMNGSYTKLYGEFLILCTY
                     IIRFYRLALTEBOQLMEKLTLYCDSTYOLIPISFVIGFYTLVTPMNMQENLEPM
                     GENTPAHEHQLEKLSLPHNMFVPMVWPMANLAKMAGLRIDPLILSLNENNTLR
                     TCGGLAYDWISIPLYTQVYVAVVSPELTLGVROFLNPAAYRHEHDLVYPR
                     DMVMNKPPEPPYTAASQFRASFGSTFNISLNKEMEQPODEEDHAGILGR
                     FLGLOSHDHPPRANSRTKLMPKRRESLHGLPKNHKAQONRGQDNKAMTKAY
                     DAFKSNAPLYKRGGITSAPOTLSPTPMFFLEPSAPSKLHSVTGIDTRKDSIKTVSSG
                     AKSFELSESQGLMEHPVSOVARKTVEENLIDMPEIPENHLKEPLDEOSPNTNHTT
                     LKDHDPYVALENDEANS"
BASE COUNT      540 a      638 c      523 g      470 t
ORIGIN
Alignment Scores:
Pred. No.:      5,36e-179
Score:          1910.00
Percent Similarity: 74.32%
Best Local Similarity: 64.19%
Query Match:    64.88%
DB:              9
                Gaps:          10
US-09-622-964-29 (1-551) x AF073501 (1-2171)
QY      1  MethThierThrAsnLysValAlaAsnAlaArgLeuGlySerPheSerLeu 20
DB      89  ATGACCTACACTTACACAAAGCGCTAATAGCCGGCTTAGCTCTTCCGCCCTG 148
QY      21  LeuLeuGlyTrpArgGlySerIleTyrLysLeuLeuTyrGlyLeuPheValPheIle 40
DB      149  CTGCTGTGCTGGCGGACACATCTACAAAGCTGCTATAGGAGATTCTTATCTTCTG 208

QY      41  PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerPheGlnGlnLeu 60
DB      209  CTGCTGCTACTACATCATCCCGCTTATTATTATAGCTGCTGCCTCAGGAGAACACACAGCTG 268
QY      61  LeuPheGlnLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
DB      269  ATGTTTGAGAAACTGACTCTGTATTCTGCACACTACATCCACTCATCCCTATTCCTTC 328
QY      81  ValLeuGlyPheTyrValIleIleValAlaSerArgTyrTrpSerGlnTyrGlnAsnLeu 100
DB      329  GTCTGTGGCTTCTACCTGACGCTGTCGCGGACCCCTGCTGGAACCACTAGACAGACTG 368
QY      101  ProTyrProAspArgLeuMetIleGlnValSerSerPheValGlnGlyLysAspGlnGlu 120
DB      389  CCGTGGCCCGACCGCTCATGAGCTGTGCTGCGCTTGTGCGAAGGACGACGACGACAA 448
QY      121  GlyArgLeuLeuAlaArgThrIleLeuIleArgTyrAlaIleLeuGlyGlnValLeuIleLeu 140
DB      449  GGCCTGCTGCTGCTGCGGACGCTCATCCGCTACGCCCAACCTGGCGACACGTCCTCATCTG 508
QY      141  ArgSerIleSerThrSerValTyrLysArgPheProThrIleuHisLeuValLeuAla 160
DB      509  CGCACCTGCACACCGCAGCTCTACAGCGCTTCCCGACCGCCACGACCTGTGTGCAAGCA 568
QY      161  GlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuGlyLeuProHisAsnThr 180
DB      569  GGCCTTATGACTCCCGGACGACACACACAGCTTGAGAACTGAGAGCTTACACACACACTG 628
QY      181  PheTrpValProThrValIlePheAlaAsnLeuSerMetLysValAlaTyrLeuGlyArg 200
DB      629  TTCGTGGTGGCTGCTGCTGCTGCTTCCCACTCATGAAAGCGTGTGAGAGTGA 668
QY      201  IleArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgTrpGln 220
DB      689  ATCCGGGACCTTATCTGCTCTCCAGACCTGCTGTAAGAGATACACACTTGTGCTACTG 748
QY      221  CysGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValAla 240
DB      749  TGTGGACACTGTATCTCTACGACGCTGATATGCCCTGTGTATACACAGCTGTG 808
QY      241  ThrValAlaValLysSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnPro 260
DB      809  ACGTGTGGCGGTGACAGCTTCTTCTGACTGTCTAGTGTGGCGGACACTTCTGACACCA 868
QY      261  AsnLysAspTyrProGlyHisGlnMetAspLeuValAlaProValPheThrIleLeuGln 280
DB      869  GCCAAGCGCTACCTGGCCATGAGCTGACCTCGTTGCGCTTCTACCTTCTCTGACG 928
QY      281  PheLeuPheTyrMetGlyTyrLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGlu 300
DB      929  TTCTTCTTCTATGTTGCTGCTGTAAGGTGCAGACGACGACATCAACACCCCTTGGAGAG 988
QY      301  AspAspAspAspPheGlnThrAsnTrpIleIleAspArgAsnLeuGlnValSerLeuLeu 320
DB      989  GATGATGATGATTTTGAACACCACTGATGTCTGACAGGATTTGACAGGTCTCTGTTG 1048
QY      321  SerValAspGlyMetHisGlnAsnLeuProPheMetGlnArgAspMetTyrTrpAsnGlu 340
DB      1049  GCTGTGATGATGATGACACGAGACCTTCTGATGAGCGGACGACATGCTGATGATAG 1108
QY      341  AlaAlaProGlnProProTyrThrAlaAlaSerAlaArgSerArgArgHisSerPheMet 360
DB      1109  CCCGAGCCACAGCCCGCTTACACAGCTTCCGCCACGATGCTGCTGACCTCTTTTATG 1168
QY      361  GlySerThrPheAsnIleSerLeuLysGlnAspLeuGlnLeuTrpSerLysGlnGlu 380
DB      1169  GGCTCCACCTTCAACATCAAGCTGAAACAAAGAGAGATGAGATTCACAGCCATACAGAG 1228
QY      381  AlaAspThrAspLysLysGlnSerGlyTyrSerSerThrIleGlyCysPheLeuGlyLeu 400
DB      1229  -----GACAGAGAGATGCTACGCTGACATCATATGCGCGCTTCTCTAGGCGCTG 1276
QY      401  GlnProLysAsnTyrHisLeuProLeuLysAspLeuLysThrLysLeuLeuCysSerLys 420

```

```

Db      1277 CAGTCCCATGATACACATCTCTCCAGGCGCAAACTCAAGCAACCACTACTGTGGCCCAAG 1336
Oy      421 AsnProleuLeu-----GLUGLYGlnCysLysAsp-----AlaasnGln 433
Db      1337 AGGGAATCCCTTCTCCACAGAGGGGCTCCCAAAAACCAAGCAGCCAAACAGAACGTT 1396
Oy      434 LysasnGlnLysAsp-----ValTTPlysPheLeuLysCysVal 451
Db      1397 AGGGCCAGGAAGACCAAGAGCCCTGGAGAGCTTAAGGCTGTGGAGCCCTTCAAGTCTGCC 1456
Oy      452 ProArpPheLysArGlySerHisCysGlyProGlnAlaProSerSer----- 468
Db      1457 CCACGTATCAGAGGCGACGCTACTACTGCTCCCAACAGACACCCCTCAGCCCACTGCC 1516
Oy      469 -----HisProthrLysLysSerHisLysCysGlyProGlnAlaProSerSer----- 482
Db      1517 ATCTTCTTCCCTCCCTAAGACATCAGCGCGCTCAAAAGCTTCACAGTGTCAAGGCAATAGAC 1576
Oy      482 ----- 482
Db      1577 ACCAAAGACAAAGCTTAAGACTGTAGTCTGTGGGCCAAGAAAAGTTTGAATGTGCTC 1636
Oy      483 -----AspGlyProSerThrAspTyrGlnGlnLysCysHisMetLysLysThr 499
Db      1637 TCAGAGAGCGATGGGCGCTTGATGAGACACCCAGAAATCTCAAGTAGAGGAGAAACT 1696
Oy      500 ValGluPheAsnLeu-----AsnIleProGlnSerProthrGlnHisLeuGlnArgArg 518
Db      1697 GTGGAGTTTAACCTCAGCGATATGACAGAGATCCCGCAAAATCCTCAAGAA-----CCT 1753
Oy      519 LeuAspGlnMetSerThrAsnIleGlnAlaLeuMetLysGlnHisLysLysTyr---- 537
Db      1754 TTGGACAAATCCACCAACATATACACACTACACTCAAGATCAATGATCTTATTTGG 1813
Oy      538 -----ProTyrArgAspGlnAlaGlyThrLysPro 547
Db      1814 GCCTTGGAAAACAGGATGAGACATCTCTAACCCT 1849

RESULT 2
AF057169
LOCUS
DEFINITION Homo sapiens bestrophin (VMD2) mRNA, linear PRI 17-OCT-1998
complete cds.
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2210)
Petukhin, K., Koist, M.J., Bakall, B., Li, W., Xie, G., Marknell, T.,
Bergren, A.A., Forsman, K., Holmgren, G., Andarsson, S., Vujic, M.,
Metzker, M.L., Caskey, C.T. and Wadell, C.
Identification of the gene responsible for Best macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
9662395
2 (bases 1 to 2210)
Petukhin, K.
Direct Submission
Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA
FEATURES
source
1..2210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2210
/gene="VMD2"

```

```

CDS
105..1862
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
spliced isoform"
/codon_start=1
/product="bestrophin"
/protein_id="AAC6433.1"
/db_xref="GI:3335159"
/translation="MTTYSQVANAALGSPFSLNLCMRGSIYKLYGELFLICY
IIRFTRALTEBOQLMEFLTYCDYSLQIDLEISVLEGEYTYVTRRMNGYENLPH
PDRMLSLVSGFVEGKDEQGRLLRLIRYANGNVILRSSTAYAKRPSAOHLVOA
GFMTPAEHKLQLESLPHNMFVPMVAFANLKMKNLGRIDPILLOSLEMMTLR
TQCHLYAVDMSILPLYTVYTVAVYSEFLICVGRQPLNFAKVPGEHDLVVPF
DYLOFFEYVGLKVAEOLINPGEDEDDDEFTWIMYDRNLQVSLAVDEHODLPHEP
DMYWNKDEPPPYTAASAPFRASRSGSTFNLSLKEEMEPNOEDEDAAGTIGR
FLGLOSDDHPPRANSRKLMLPKRRESLHEGLPNHKAQNVGORDNKAATKAV
DAFSAPLYORPGYVSAPOPLSPMPFEPLEPLPSKLSHVSYTGIDTKSLKYSSG
AKSFELSESDGALMEHPEVSQVRRKTVFELTDMPEIPENHLEKPELOSPTNIHTT
LKDHPDYPVALENRDEAHS"
BASE COUNT 556 a 647 c 531 g 476 t
ORIGIN

Alignment Scores:
Pred. No.: 5,5e-179 Length: 2210
Score: 1910.00 Matches: 380
Percent Similarity: 74.32% Conservative: 60
Best Local Similarity: 64.19% Mismatches: 102
Query Match: 64.88% Indels: 50
DB: 9 Gaps: 10

US-09-622-964-29 (1-551) x AF057169 (1-2210)
Oy      1 MethTrIleThrTyrThrAsnLysValAlaAsnAlaArgLeuGlySerPheSerSerLeu 20
Db      105 ATGACATCACTTACACAGCAAGCTGCTAATAGCCCGCTTACTCTCCCGCTG 164
Oy      21 LeuLeuGlyStrParGlySerIleTyrLysLeuLeuTyrGlyLysPheLeuValPheIle 40
Db      165 CTGCTGTGCTGGCGGGGACACATCTACAGCTCTATAGCGCAGTCTTAATCTTCCTG 224
Oy      41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGlnLeu 60
Db      225 CTCTGTACTACATCATCTCCCTTATTATAGCGTGGCCCTCAGGAGACAGACAGCTG 284
Oy      61 LeuPheGluLysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
Db      285 ATGTTTGAGAAACTGCTGTATTTGCGACACTACATCCAGCTCATCCCATTTCTCTC 344
Oy      81 ValLeuGlyPheTyrValThrLeuValAlaSerArgTPTPTSerGlnTyrGlnAsnLeu 100
Db      345 GTGCTGGGCTTCACGTACGCTGCTCGTACCCCTGCTGGAGACCACTACAGAACTG 404
Oy      101 ProTTPProAspArgLeuMetIleGlnValSerSerPheValGluGlyLysAspGlnGln 120
Db      405 CCGTGGCCCGACCGCTCATAGCCTGTGCGGCTCGTGGAGCAAGCAAGCAGACCA 464
Oy      121 GlyArgLeuLeuArgArgThrLeuIleArgTyrAlaIleLeuGlnGlnValLeuIleLeu 140
Db      465 GCGCCGCTGTGGCGGCGAGCTCATCCGCTACGCCAACCCTGGGCAACGCTCATCTCTG 524
Oy      141 ArgSerIleSerThrSerValTyrLysArgPheProThrLeuHisLysLeuValLeuAla 160
Db      525 CGCAGCGTCAGCAGCGAGCTCAACAGCGCTCCCGCCAGCCCGCAGCAACCTGTGCAAGC 584
Oy      161 GlyPheMetThrHisGlyGlnHisLysGlnLeuGlnLysLeuLysLeuProHisAsnThr 180
Db      585 GGCCTTATGACTCCGCGAGAACACAAAGCACTGTGAGAAACTGAGCCCTACACACACANG 644
Oy      181 PheTyrPAlaProTyrPAlaTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlnGlyArg 200
Db      645 TTTGGGTGCGCTGTGGTGTGGTTCGCAACCTGTCAATAAGCGCGCTTGGAGGTCA 704
Oy      201 IleArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgThrGln 220

```

```

Db      705 ATCCGGAGACCTTACCTGCTCCAGACCTGCTGTAAGAGATGAACCTTGGCTACTG 764
Qy      221 CysGlyGlnLeuYrAlaYrAspTrpIleSerIleProLeuValYrThrGlnAlaVal 240
Db      765 TGTGGACACCTGTATCTCCACGACTGATAGATCCACTGCTATACACAGGTGGT 824
Qy      241 ThrValAlaValYrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAspPro 260
Db      825 ACTGTGGCGGTGTACAGCTTCTTCTGACTTGTCTAGTGTGGCGGACAGTTCTGAACCA 884
Qy      261 AsnLysAspYrProGlyHisGlnMetAspLeuValProValPheThrIleLeuGln 280
Db      885 GCCAAGGCTTACCCTGGCCATGAGCTGACCTGTCGTCGTCGTCCTTACCTTCCCTGAG 944
Qy      281 PheLeuPheYrMetGlyTyrPheLysValAlaGlnGlnLeuIleAsnProPheGlyGlu 300
Db      945 TTCTTCTTCTATGTTGGCTGCTGCTGTAAGGTGGCAGAGCTCATCAACCTTGGAGAG 1004
Qy      301 AspAspAspAspPheGlnThrAsnTrpIleLeuAspArgAsnLeuGlnValSerLeuLeu 320
Db      1005 GATGATGATGATTTGGACACCACTGATGTCGACAGGATTTGCGAGGTGTCCTGTTG 1064
Qy      321 SerValAspGlyMetHisGlnAsnLeuProProMetGlnArgAspMetYrTrpAsnGlu 340
Db      1065 GCTGTGATGATGATGACACAGAGACCTGCTGCTGATGAGCCGACATGACTGGAATAG 1124
Qy      341 AlaAlaProGlnProProTyrTrpAlaAlaSerAlaArgSerArgHisSerPheMet 360
Db      1125 CCCAGGCCACAGCCCCCTACACAGCTGCTCCGCCAGTTCCTGACGCTCTCTTATG 1184
Qy      361 GlySerThrPheAsnIleSerLeuLysGlnAspLeuGlnLeuTrpSerLysGlnGlu 380
Db      1185 GGCTCCACCTTCAACATCAGCTCTGACACAGAGAGAGATGAGATTCACGCCCATCGAG 1244
Qy      381 AlaAspThrAspLysLysGlnSerGlyYrSerSerThrIleGlyCysPheLeuGlnGlu 400
Db      1245 -----GACGAGAGAGATGCTCACGCTGGCATCATGCGCTTCCCTAGGCTG 1292
Qy      401 GlnProLysAsnTrpHisLeuProLeuLysAspLeuLysThrLysLeuLysSerLys 420
Db      1293 CAGTCCCATGATCACAATCCTCTCCAGGCAAACTCAAGAGACCAAACTACTGTGGCCAA 1352
Qy      421 AspProLeuLeu-----GlnGlyGlnCysLysAsp-----AlaAsnGln 433
Db      1353 AGGGAATCCCTCTCCACGAGGCGCTGCCAAAACACACAGCCAAACAGAACGTT 1412
Qy      434 LysAsnGlnLysAsp-----ValTrpLysPheLysGlyLeuAspPheLeuLysCysVal 451
Db      1413 AGGGCCAGAGAACACAAAGAGCCCTGGAAGCTTAAAGCTGTGGAGCCCTTCAGATGCGC 1472
Qy      452 ProArgPheLysArgArgGlySerHisCysGlyProGlnAlaProSerSer----- 468
Db      1473 CCACCTGATCAGAGCCAGGCTACTACAGTGGCCACAGAGAGCCCTTCAGCCCACTGCC 1532
Qy      469 -----HisProThrGlnGlnSerLysAspSerSer-----SerAspThrGly 482
Db      1533 ATGTCTCTCCCTTGAACCATCAGCGCCGTCAAAAGCTTCAAGTGTCAAGGATAGAC 1592
Qy      482 ----- 482
Db      1593 ACCAAAGCAAAAGCTTAAAGCTGTGAGTCTGGGGCCAAAGAAATTTTGAATTCCTC 1652
Qy      483 -----AspGlyProSerThrAspYrGlnGlnIleCysHisMetLysLysLysThr 499
Db      1653 TCAGAGAGCATGGGCGCTTGATGGAGCCACCAAGATATCTCAAGGAGAGGAAACGT 1712
Qy      500 ValGlnPheAsnLeu-----AsnIleProGlnSerProThrGlnHisLeuGlnAlaArg 518
Db      1713 GTGGAGTTTAACTGAGGATATGCGCAGATCCCGAAATACCTCAAGAA-----CCT 1769
Qy      519 LeuAspGlnMetSerThrAsnIleGlnAlaLeuMetLysGlnHisLysLeuSerYr----- 537

```

```

Db      1770 TTGACAAATACCAACCAACATACACTACACTCAAGATCATGATCCTATTGCG 1829
Qy      538 -----ProTyrArgAspGlnAlaGlyThrLysPro 547
Db      1830 GCCTTGAAACAGGAGATGAGCATCTCTTAACCT 1865

```

```

RESULT 3
LOCUS   AF057170
DEFINITION Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product.
ACCESSION AF057170
VERSION   AF057170.1 GI:3335160
KEYWORDS
SOURCE   Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS   Petukhin, K., Koist, M.J., Bakall, B., Li, W., Xie, G., Marknell, T.,
Bergen, A.A., McGarity-Dugan, V., Figueroa, D., Austin, C.P.,
Metzker, M.D., Caskey, C.T. and Madeline, C.
TITLE      Identification of the gene responsible for Best macular dystrophy
JOURNAL    Nat. Genet. 19 (3), 241-247 (1998)
MEDLINE    98324772
PUBMED     9662395
REFERENCE 2 (bases 1 to 2420)
AUTHORS   Petukhin, K.
TITLE      Direct Submission
JOURNAL    Submitted (03-APR-1998) Human Genetics, Merck Research
LABORATORIES, West Point, PA 19486, USA
FEATURES
source
1..2420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2420
/gene="VMD2"
105..1412
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
spliced isoform; contains exon 7."
/codon_start=1
/product="bestrophin"
/protein_id="AAC64344.1"
/db_xref="GI:3335161"
/translation="MTITYSQVNAARLGSFRLICRGSTYKILNGEPLILCY
PRMLSVSGFEGKDEQGRLLRLIRAILGNVLLIRSVTVAVTRMVAQENLPA
GMPTRAEKOLEKSLPRINKEVFWPWFANLSMKAWGLRDRILLOSILNENKTLR
TTLQFFTVGMLKGLSRALGLWRHGGHGOQLLETRMOCQERKVRVRSQAMMT
PVTPATRAEAGESLEPGRRLIMQSSSTLEMMIMILRPTGLSTGICRCPCMLNR
CTRILGMSRCTGSPSPSTOLLPSSVPEPLMAPPTSA"
BASE COUNT 605 a 695 c 612 g 508 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 3 77e-172 Length: 2420
Score: 1841.50 Matches: 379
Percent Similarity: 66.52% Conservative: 60
Best Local Similarity: 57.42% Mismatches: 103
Query Match: 62.55% Indels: 119
DB: 9 Gaps: 11

```

US-09-622-964-29 (1-551) x AF057170 (1-2420)

```

Qy      1 MethrIleThrYrThrAsnLysValAlaAsnAlaArgLeuGlySerPheSerLeu 20
Db      105 ATGACCATCACTTACACAGCAAGGTGGCTTAATGGCCGCTAGGCTTCTTCCGCCCTG 164

```


OY 21 LeuLeuCySerPArgGlySerIleTyrIleuLeuTyrGlyGluPheValPheIle 40
 DB 165 CTGCTGCTGGCGGGGCGACATCTACAGAGCTGATATGCGAGTTCTTATCTTCG 224
 OY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgTyrMetValLeuSerSerArgGlnLeu 60
 DB 225 CTCTGCTACTACATCATCGCTTATATATAGCTGGCGCTACAGGAAACACACAGCTG 284
 OY 61 LeuPheGlnIleuValLeuTyrCysArgSerTyrIleGlnIleuIleProIleSerPhe 80
 DB 285 ATGTTTGAAACAGTACTCTGTATGCGACAGTACATCCAGGTCATCCCATTCCTTC 344
 OY 81 ValLeuGlyPheTyrValIlePheValValSerArgTyrPheSerGlnTyrGlnLeu 100
 DB 345 GTCTGGCTCTTCTACGTAGCGCTGTGTCAGCCCGCTGGAGAACAGTACAGAGACTG 404
 OY 101 ProTyrProAspArgLeuMetIleGlnValSerSerPheValGluGlyAspGlnGlu 120
 DB 405 CCGTGGCCCGACCGCTCATAGACCTGTGTCTCGGCTGTGTCGAGGAAAGACAGAC 464
 OY 121 GlyArgLeuLeuArgTyrIleuIleArgTyrAlaIleLeuGlyGlnValLeuLeu 140
 DB 465 GGGCGGCTGTGGCGGCGACGCTACCGCTACCGACCTGGGACGCTGCTCATCTG 524
 OY 141 ArgSerIleSerThrSerValTyrTyrAspArgPheProIleuIleValLeuAla 160
 DB 525 CCGAGGCTACGACCGACAGTACAGAGCTTCCCGAGGCGCCACAGCTGCTGACAC 584
 OY 161 GlyPheMetThrHisGlyIleuHisArgGlnLeuGlnIleuGlyLeuProHisAsnThr 180
 DB 585 GCGTTTATGACTCCGCGACAGACACAGAGCTGGAGAAATGAGCTTACACACACATG 644
 OY 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetIleValTyrGlnGlyArg 200
 DB 645 TTCTGGTGGCTGTGGCTGTGTGCTGCTGCAACCTGCAAGAGGCTGGCTGGAGTGA 704
 OY 201 IleArgAspThrValLeuLeuGlnSerLeuMetAsnGluValCysThrLeuArgThrGln 220
 DB 705 ATCCGCGACCTATACCTGCTGTCAGAGCTGTGAGAGATGAAACACTTGGCTGACTCAG 764
 OY 221 CysGlnIleuLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnVal 240
 DB 765 TGTGGACACCTGTATGCTTACGACTGATATGATCCAGTGTGTATACACAGGTGTG 824
 OY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnPro 260
 DB 825 ACTGTGGCGGTGTACACTTCTTCTGACTGTGTGTAGTTGGCGGAGTTTCTGAAACCA 884
 OY 261 AsnLysAspTyrProGlnHisGlnMetAspLeuValValProValPheThrIleLeuGln 280
 DB 885 GCCAAGGCTACCTGCGCCATGAGCTGAGACCTGTGTCGCGTCTTACGTTCTCTGAG 944
 OY 281 PheLeuPheTyrMetGlyTyrPheLeuVal 290
 DB 945 TTCTTCTTATGTGTGCTGTGCTGAAAGT-GGGCTCTTCCAGGCGCTGTGGCTGAG 1003
 OY 290 290
 DB 1004 GCATGGCCAGAGGGGTATGCGCACAGCTGCTGAGAGAGATGCAAGTCTCAGAAAG 1063
 OY 290 290
 DB 1064 GAAGGTCTCAGCGGTAGAAAGACGACGAGCGGTGTGGCGACACCTGTAAATCCAGCTAC 1123
 OY 291 291
 DB 1124 TCGGAGGCTGAGCGAGAGATGCTTGAACCGCGGAGCGGAGGTGGTGGCGAGAG 1183
 OY 293 GlnLeuIleAsnProPheGlyGlyAspAspAspPheGlnTyrAsnTyrIleIleAsp 312
 DB 1184 CAGCTCATCAACCCCTTGGAGAGATGATGATATTTTGAACCACTGATGTGTGCAGC 1243
 OY 313 ArgAsnLeuGlnValSerLeuLeuSerValAspGlyMetHisGlnAsnLeuProPromet 332

DB 1244 AGGAATTTCCAGAGTCTCCCTGTGGCTGTGATGAGATGACACAGAGACTCTCGGATG 1303
 OY 333 GluArgAspMetTyrTyrAsnGlnValAlaProGlnProProTyrThrAlaAlaSerAla 352
 DB 1304 GAGCGCGAGATGTACTGTGAATATGAGCCGAGACAGCCCTTACAGAGCTGTCCGCC 1363
 OY 353 ArgSerArgTyrHisSerPheMetGlySerThrPheAsnIleSerLeuTyrGlyAsp 372
 DB 1364 CAGTTCCCTCCAGCTCTTATGTGGCTTCCACCTTCAACATCAGCTGACCAAGAGAG 1423
 OY 373 LeuGluLeuTyrSerTyrGlyGluAlaAspThrAspTyrArgSerGlyTyrSerSer 392
 DB 1424 ATGAGAGTCCAGCCCAATCAGAG- -GACAGAGAGATGCTCTACGCTGCC 1471
 OY 393 ThrIleGlyCysPheLeuGlnIleGlnProLysAsnTyrHisLeuProLeuLysAspLeu 412
 DB 1472 ATCATTTGGCGCTTCTAGGCTGTGAGTCCCATGATCACCATCTCTCCAGGGCAACTCA 1531
 OY 413 LysThrLysLeuLeuCysSerLysAsnProLeuLeu- -GluGlyGlnCysLysAsp 430
 DB 1532 AGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCAGAGAGGCTGCCCAAAAC 1591
 OY 431 431
 DB 1592 CACAGAGCAACCCAAACAGACCTTATAGGGGCGCAGAAACACACAGCGCTGAACTTAAG 1651
 OY 444 GlyLeuAspPheLeuLysCysValProArgPheLysArgGlySerHisCysGlyPro 463
 DB 1652 GCTGTGAGCGCTTCAAGTCTGCCCATGTATCAGAGCGGCTTACTAGTGCCTCA 1711
 OY 464 GlnAlaProSerSer- -HisProThrGlnIleSerAlaProSerSer 478
 DB 1712 CAGAGCGCCCTTACGCCCTCCATGCTTCTTCCCTTCAACATGACAGCGCTCAAG 1771
 OY 479 479
 DB 1772 CTTCACAGTGTACAGGATGAGACACCAAAAGCTTAAAGACTGTGACTTCTGGG 1831
 OY 483 483
 DB 1832 GCCAAGAAATTTGAATTCCTCTCAGAGAGCGATGGGCTGTATGAGAGACCCAGAA 1891
 OY 492 IleCysHisMetLysLysThrValGluPheAsnLeu- -AsnIleProGlnSerPro 510
 DB 1892 GTATCTCAAGTGAAGAGAGAAACCTGTGAATTTTACCTGACGATATGCCAGATGCC 1951
 OY 511 ThrGluHisLeuGlnIleArgTyrLeuAspGlnMetSerThrAsnIleGlnAlaLeuMet 530
 DB 1952 GAAATTCACCTCAAGAA- -CCTTGGAAATATCACCAACCAACTATACACTACACTC 2008
 OY 531 LysGluHisAlaGluSerTyr- -ProTyrArgAspGlnAlaGlyThrLysPro 547
 DB 2009 AAAGATCACATGATCTTATTTGGGCTTGGAAACAGGAGATGAAACATCTTCAACT 2068
 RESULT 4
 AK096459
 LOCUS
 DEFINITION
 AK096459
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
 Arita,M., Mutsashino,K., Yuki,H., Hara,H., Sugiyama,T., Irie,R.,
 Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
 Kawai,Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES

Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Magatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Nagahari, K., Masuhio, Y., Nagai, K. and
Isogai, T.
NEO human cDNA sequencing project
Unpublished
2 (bases 1 to 2500)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, Fty Project (HRI Team): 2-6-7
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB) (supported by Japan
Key Technology Center etc.); 5'-63'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2500
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NTONG2009233"
/tissue="tongue"
/clone_id="NTONG2"
/note="cloning vector: pME18SFL3"
229..2233
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04797.1"
/db_xref="GI:21755963"
/translation="MTWYSSSKYANATFFGFRHLKMGSTYKLLREFYFAVLV
AISLYVRLITGVOKRTEKLTICDVAEIQPVFVLFVTLVNNMMNQFNVA
PDLMLPSSVSGDEHGLRRLMRYVNLSTLIFSVETAYKKRPDHYEA
GKPTDERKLNHKLSPHLKWPFIWFNQLATKANEGRIDSDLOSIMENRIR
SMCSLEFGIDWVIGIPLYQVYVLAIVFFACILGROVDPTKRYAGHDLDLTPF
TLQFFYAGMLKVAEQLINFGEDDDDEFTWICIDRINQVSLAVDMHSLPMK
DLYWDSAPRPYTLAAADYCIPIFSGISVOMGICSGSPDEMLMDYKSHSMI
RKYRRLSAHEHSPRSYRROTSDSMFLPRDLSFADLDLDPVRNPRASPW
KKSCPEGSPTLHFSMGLSTIRTSOTSLDLPQSSVRSPIKMLPEVEVLTAA
EAVPVSQGHDSATLISSEFTVDSKTEQDQPGSVSILSPSEKTPGSGSPOT
VSAEAENFENCEDEPDFTLRMSLPFLSSHTSLGNLSLSPDSQSPALLIDTETS
SEISGINTYAGSVSDMLYLMENLDTEETDIIEINKETESPK"

CDS

BASE COUNT 689 a 632 c 563 g 616 t
ORIGIN

Alignment Scores:

Score: 4.93e-117 Length: 2500
Percent Similarity: 1285.00 Matches: 276
Best Local Similarity: 63.15% Conservative: 77
Query Match: 49.37% Mismatches: 137
43.65% Indels: 69
Gaps: 9

US-09-622-964-29 (1-551) x AK096459 (1-2500)

QY 1 Methrilethrythrinalysvalalalasnalaargleuglyserpheserleu 20
DB 229 ATGACTGTCACTTCTCCAGTAAAGTAGCAACTTTTGGATTTCATAGCTTA 288
QY 21 Leuencysttrparglyseriletyrlyleuleuletyrlyglunpheuleu 40
DB 289 CTCTCAAGTAGAGGACGACATCTCAAACTCTGACAGGGAATTATGTTTGGCT 348
QY 41 Pheleutyrtyserileargglyleutyrtargmetvalleuserserapcginleu 60
DB 349 GTCTTTATACACCAATTAAGTTGGTACAGATTGTTACTACAGAGACTCCAAAAGCT 408
QY 61 Leupheuglyleuleuleutyrcysaspseryrilegileuleileproilaserphe 80
DB 409 TACTTTGAAATAATATCATCTTACTGTGACAGATATGCTGAACAATTCAGTAACCTT 468

QY 81 ValleuglypheyvalaThrleuvalvalaseratgtrpserglntrygluanleu 100
DB 469 GTGCTTGCGTTTATGTTACTGTGGTAGGAACGATGGTGGAAACGATTGTGAATTGG 528
QY 101 ProtrpProasparglymetileginalyserPheValGlnGlyasapcglu 120
DB 529 CCTGGCCAGACGCGTAATGTTCCATCTCTACAGAGTTCACGGAAGCAGCAGCAGC 588
QY 121 GlyArgLeuLeuArgThrleuLeuLeuArgTyrAlaIleLeuGlyValleuLeu 140
DB 589 GGGCGGCTCTTGAAGAGAGCGGTGATGGCTACGCAATCCACATGACACGCTGTCATCTT 648
QY 141 ArgserIleSerThrserValtyrlysaArgPheProThrleuHisIleValleuAla 160
DB 649 CCGTCGTGACGACGCTGTGTACAAAGATTCCACATGACACGCTGTTGAAGCA 708
QY 161 GlyPheMetThrHisGlyGlnHisIleGlnLeuGlnLeuGlyLeuProHisAsnThr 180
DB 709 GGTTTTATGACACAGATGAAAGAAATATTCACACCCATGCTCCTCATCTGAA 768
QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrleuGlyArg 200
DB 769 TATGGGTTCCATTCATCTGTGTTGAATCTTCACATTAAGCCGGAATGAAGTAGCA 828
QY 201 IleArgAspThrValleuGlnSerLeuMetAsnGlnValCysThrleuArgThrGln 220
DB 829 ATCAGAGACACTGTGATGTCGAATCATTCATGACGAAATGATGATACCGCTCTGG 888
QY 221 CysGlyLeuLeuTyrAlaTyrAspTrpIleSerIleProleuValTyrThrGlnVal 240
DB 889 TGCAGCGCTTATCGGTTATGACGTGGTGGATTCGCCCTGTTACCCAGGTGTGC 948
QY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnPro 260
DB 949 ACCTTGCTCTCATACCTCTCTCTGCTGCTGATGGACGCGCAGTTTGGATGCC 1008
QY 261 AsnLysAspTyrProGlnHisGlnMetAspLeuValAlaProValPheThrIleLeuGln 280
DB 1009 ACCAAAGGCTGACGAGGCACTGATGATCTTTCATTCCTCATTCACCTCTCTCA 1068
QY 281 PheLeuPheTyrMetGlyTyrPheLysValAlaGlnLeuIleAsnProPheGlyGln 300
DB 1069 TTCTTCTTATGACAGATGCTTAAGTACGACGACGCTTATCAACCCCTTTGGAA 1128
QY 301 AspAspAspPheGlnThrAsnTrpIleIleAspArgAsnLeuGlnValSerLeuLeu 320
DB 1129 GATGATGATGATTTGAAACTAAGTGCATTCACAGAAATTTGACAGTCTCTTTTA 1188
QY 321 SerValAspGlyMetHisGlnAsnLeuProPheMetGlnArgAspMetTyrTrpAsnGln 340
DB 1189 GCTGGAGCAAGATGACATGAGCTTACCCAAAGTAGAAGACATTCATGCGAGCAT 1248
QY 341 AlaAlaProGlnProPheTyrThrAlaAlaSerAlaArgSerArgTrpHisSerPheMet 360
DB 1249 TCTGCTGCTGCCACCATACATTCGACGTCGTCACACTGCATACCCCTCATTTG 1308
QY 361 GlySerThrPheAsnIleSerLeuLysGlnAspLeuGlnLeuTrpSerLysGlnGln 380
DB 1309 GGGTCACACATCCAGATGGGCTGTGCTGCTCGAC-----TTTCTGACGAGGAG 1359
QY 381 AlaAspThrAspLysLysGlnSerGlyTyrSerSerThrIleGlyCysPheLeuGln 400
DB 1360 TGGCTGTGATATATAGAAAGCATGCGCATCG----- 1392
QY 401 GlnProLysAsnTrpHisLeuProleuLysAspLeuLysTrpLysLeuLeuCysSerLys 420
DB 1393 -----CATTCATGATATAGAAAGCAAGCGGTCTCTGAGGCCACAGAA 1437
QY 421 AsnProLeuLeuGlnGlyGlnCysLysAspAlaAsnGlnLysAsnGlnLysAspValTrp 440
DB 1438 CAC----- 1440

[illegible]

US-09-622-964-29 (1-551) x BC019528 (1-1957)

```

OY 1 MetThrIleThrTrpAsnLysValAlaAsnAlaArgLeuGlySerPheSerIleu 20
DB 60 ATGACGCGCTACCTACAGCAGAGTGGCGAATGCCCGTGGCTTCTCGACAGCTG 119
OY 21 LeuLeuGlySerTrpArgIleSerIleTyrLysLeuLeuTyrGlyLysLeuValPheIle 40
DB 120 CTGCTGCTGGCGCGGAGACATCTACAGCTCTGTGGCGAGAGCTGTATGTTCCG 179
OY 41 PheLeuTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspIleuLeu 60
DB 180 GAGCTCTACCTGCTAGACCGCCCTATCGCTCTTCTTCTGCGAAGAGAGAGAGCGC 239
OY 61 LeuPheGlyLysLeuAlaLeuTyrCysAspSerTyrIleGlnIleProIleSerPhe 80
DB 240 TACTCTGAGAGAGCTGTATATACGCGAGCAGTACGCCAGCTCATCTCCGCTCTTC 299
OY 81 ValLeuGlyPheTyrValThrLeuValValSerArgTyrPyrSerGlnTyrGlnLeu 100
DB 300 GTACTTGGCTTCTACGTACTCTGTGGTGCATGCTGTGGAAACGATCATATGCTATG 359
OY 101 ProTyrProAspArgLeuMetIleGlnValSerSerPheValGlyLysAspGlyLys 120
DB 360 CCTGTCCGAGCGACTCATATGCTAGTGTGGTGGCACCGTGCATGGGCGAGAGATCGA 419
OY 121 GlyArgLeuLeuAlaGlyThrLeuIleArgTyrAlaIleLeuGlyLysValIleu 140
DB 420 GCGCCCTCTACCGCGCGACGCTCATGCTGCTACGACGAGGCTCTCCGCGGTGATCTT 479
OY 141 ArgSerIleSerThrSerValTyrLysArgPheProThrLeuHisLeuValIleu 160
DB 480 CGTCTGTGACAGAGAGCTCTTCAAGGCTTCCCACTATACACACGCTGGCGAGGCT 539
OY 161 GlyPheMetThrHisGlyLysLysGlnLeuGlyLysLeuGlyLysLeuProHisLeu 180
DB 540 GGAATTATGACCGGAGAGAGCGCAAGATTGAGAACTGAATTCGCTCTACAAACA 599
OY 181 PheTrpValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyLys 200
DB 600 TACTGGTGGCTGCTATGCTTCTGACCTGTGGACGCGCGAGCGCGGAGAGGCGGC 659
OY 201 IleArgAspThrValLeuLeuGlnSerLeuMetLeuGlnValCysThrLeuArgThrGln 220
DB 660 ATCCGCGAGACAGAGCTCCCTAAAGTGTCTACTAGAGAGCTGAATGCTTTCCGAGCA 719
OY 221 CysGlyGlnLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnVal 240
DB 720 TGTGGATGCTGTCTACAGACGATGATGATGATGATGATGATGATGATGATGATG 779
OY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnPro 260
DB 780 ACTATCGACAGTGTACGCTACTTCTGTGCTGCTCATCGGCGCTGCTCATGACCC 839
OY 261 AsnLysAspTyrProGlyHisGlnMetAspLeuValValProValPheThrIleLeuGln 280
DB 840 GCACAGGCGCTCAAAAGACACACCTGAGCTTACGCTTACCTTACCTGCTGACG 899
OY 281 PheLeuPheTyrMetGlyTyrPheLysValAlaGlnLeuLeuAsnProPheGlyLys 300
DB 900 TTCTTCTTCTTACGCTGCTGCTTAAAGTGAAGCAGACAGCTTATTAACCTTTTGA 959
OY 301 AspAspAspAspPheGlyThrAsnTrpIleIleAspArgAsnLeuGlnValSerLeu 320
DB 960 GACACACACGACTTGAACCACTTCTTATGACCGCAACTTCCAGGTGCTCATGCTA 1019
OY 321 SerValAspGlyMetHisGlnLeuLeuProMetGlnArgAspMetTyrTrpAsn 340
DB 1020 GCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
OY 341 AlaAlaProGlnProProTyrThrAlaAlaSerAla---ArgSerArgArgHisSerPhe 359
DB 1080 GCAGAGAGCTGCGCGCGCTACACCGCGCGCTTCTGCTTCTGCTTCTGCTTCTGCT 1139

```

```

OY 360 MetGlySerThrPheAsnIleSerLeuLysGlyLysLeuGlnLeuTrpSerIleu 379
DB 1140 CAGGCTCTCACCTTGTATATGAGCTGTGGTAAAGAGACATGATGATCCAGCGCTG 1199
OY 380 GlnAlaAspThrAspLysGlySerGlyTyrSerSerThrIleGlyCysPheLeu--- 398
DB 1200 GCGGTGATGCGCGCTGGGAGAG---GTTACAGCGGAGCTTCTGTCAG 1244
OY 399 -----GlyLeuGlnProLysAsnTyrHisLeuProLeuLys 410
DB 1245 CGCTCTCTCCGCGCGCGCGGCTTCTGTGGGTCCGCTGGGCGCGCGCTGCTGTTA 1304
OY 411 -AspLeuLysThrLysLeuLeuCysSerLysAsnProLeuGlnGlyLysLys 430
DB 1305 AGCGCTAAATAACA---GCTGCGTGTACAGAGGCTTCACTGCTGCACGCTGTGCA 1361
OY 430 PAlaAsnGlnLys----- 434
DB 1362 GCGCGCGCGAGATGGCGGAGCGGTGAGTGTGCGGTTGTGGAGACCATTTGTCACCTAGC 1421
OY 434 ----- 434
DB 1422 CTGCGGAGCGCTGAAGTGAAGCTCCCGCATGCTCTGAGCGCGCTGCGCCATCCG 1481
OY 435 -----AsnGlnLysAspValTyrPheLysGlyLysLeuAspPheLeuLys----- 449
DB 1482 CCAACTCTGAACTTTTACACCGTGTGATTCGCGGCGCTGCGGCTGCTGCTGTTA 1541
OY 450 ---CysValProArgPheLysArgArgGlySerHisCysGlyProGlnAlaProSer 468
DB 1542 TGGCTGCGCTACCGCTTATGAGAGAGAG---AAGAGAGTCCGCGCTGAGATCTCGG 1598
OY 468 HisLeuProGlnGlnSerLysProSerSerSerAspTrpGlyAspGlyProSerThr 488
DB 1599 AGAGACCGCTTACG---ACACCTTACCTTACCTGACCTGACAT---ACTGGCGA 1646
OY 488 PTrpGlnGlnIleCysHisMetLysLysThrValGlnPheAsnLeuAsnIleProG 508
DB 1647 CCACGAGGCTGCGACACACTAGCTAGAGCTTCA-----GA 1682
OY 508 userProThrGlnHisLeuGlnGlnArgArgLeu 519
DB 1683 AAGCGAGCTACTCATTTTGACACAGCGGAATCTC 1716

```

```

RESULT 8
AF440756 1908 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens vitelliform macular dystrophy 2-like protein 1 mRNA,
ACCESSION AF440756
VERSION AF440756.1 GI:21734839
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1908)
Marguardt, A., Stohr, H., Passmore, L. A., Kramer, F., Rivera, A. and
Weber, B. H.
Mutations in a novel gene, VMD2, encoding a protein of unknown
properties cause juvenile-onset vitelliform macular dystrophy
(best's disease)
Hum. Mol. Genet. 7 (9), 1517-1525 (1998)
MEDLINE 98367043
PUBMED 9700209
REFERENCE
2 (bases 1 to 1908)
Stohr, H., Marguardt, A., Nanda, I., Schmid, M. and Weber, B. H.
Three novel human VMD2-like genes are members of the evolutionary
highly conserved RFP-TM family
Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
MEDLINE 22027749
PUBMED 12032738

```

REFERENCE 3 (bases 1 to 1908)
 AUTHOR: Stohr, H., Marguardt, A. and Weber, B. H. F.
 TITLE: Three novel human YMD2-like protein genes are members of the
 JOURNAL: evolutionary highly conserved RFP family
 REFERENCE 4 (bases 1 to 1908)
 AUTHOR: Stohr, H., Marguardt, A. and Weber, B. H. F.
 TITLE: Direct Submission
 JOURNAL: Submitted (30-OCT-2001) Human Genetics, University of Muenzberg,
 Biozentrum, Am Hubland, Muenzberg 97074, Germany
 FEATURES
 source
 1.1908
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p13.12-p13.2"
 /note="VMD2L1; putative membrane protein; Pfam01062;
 norm.family=8"
 /codon_start=1
 /product="YMD2L1 form macular dystrophy 2-like protein 1"
 /protein_id="AA076995.1"
 /db_xref="GI:21734840"
 /translation="MTVYTPARVANAARPGFSQILLMRGSIYKLLMRLLCFGLPYM
 ALSAAYRVLEGGKRRYFKLVITCDYASLIPVSEVLGYTVLVNRMWSOYLCPL
 PDALMCVAGVGRHGRDRGLYRTLRVAGLSAVLRGTVAVRKRPTIDHYEA
 GPMTRERKRENNSSYNRYWPCWFSMLAARREGIRNSALKLLEERLVEFR
 TLDPFFYAGLKYAEQILNPFGEDEDDFTNFIIDRNQVSLAYDEMIDLVLEK
 DLYWDAARAPYATATFOLRQPSFGSPFDITLAKEDQERLDGLDPMGEAPGD
 FLORLPAGAGVYAGPLGRSLRSLRKNSGSEASTGSCAVVPEEAACGCCGD
 PLDPGLPEEAPPPAPPEPLTLIPGVEFSLVMPGRGAPAPLPSIGEENL
 A"

CDS

BASE COUNT 337 a 604 c 551 g 416 t
 ORIGIN

Alignment scores:

Pred. NO.: 1.67e-111 Length: 1908
 Score: 1227.50 Matches: 256
 Percent Similarity: 65.34 Conservative: 72
 Best Local Similarity: 51.008 Mismatches: 145
 Query Match: 41.698 Indels: 31
 DB: 9 Gaps: 5

US-09-622-964-29 (1-551) x AR440756 (1-1908)

QY 1 MetThrIleThrTyrThrasnlyValAlaAsnAlaArgLeuGlySerPheSerSerLeu 20
 DB 1 ATGACCGTCACCTACACGCGGAGTGGCGGACCGCTTCGCTGCTCCACGCTG 60
 QY 21 LeuLeuGlySerPheGlySerIleTyrLeuLeuTyrGlyIlePheLeuValPheIle 40
 DB 61 CTGCTACTGTGGCGGAGCATCTACAACTCTGTGGCGGAGCATCTCTCTCTCT 120
 QY 41 PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGlnLeu 60
 DB 121 GGGTTCATGAGCGGCGAGTGCCTACCGCTTGTCTGCTACCGAGGAGGAGCGCC 180
 QY 61 LeuPheGlyLeuValAlaLeuTyrCysAspSerTyrIleGlnLeuIlePheIleSerPhe 80
 DB 181 TACTTCGAGAGGCTGTGATTATTTGACCACTAGTACGAGCTCATCCCTGCTCTCTC 240
 QY 81 ValLeuGlyPheTyrValIleValValSerArgTyrTyrSerGlnTyrGlnLeu 100
 DB 241 GTGCTTGGCTTTATGTGAGCGTGTGTGAGAACCGCTGTGTGAGCCAGTACATGATG 300
 QY 101 ProTyrProAspArgLeuMetIleGlnValSerSerPheValGlyIleGlyAspGln 120
 DB 301 CCCCTGCGCCGACGCTCATGTGCGTGTGGCGGAGCGGACGCTCAGAGGAGCGCC 360
 QY 121 GlyArgLeuLeuArgTyrThrLeuIleArgTyrAlaIleLeuGlyGlnValLeuIleLeu 140
 DB 361 GGGCGGCTCTACCGGCGGACATCATGCGCTACGAGGAGGCTCTCGCGCTCATCTCTG 420

QY 141 ArgSerIleSerThrSerValTyrTyrArgPheProThrLeuIleHisLeuValLeuAla 160
 DB 421 GCGTCGCTCAGACACCGGCGGTGTTCAGCGCTTCCACCATACACACGCTGGAGGCT 480
 QY 161 GlyPheMetThrHisGlyGlnHisGlyGlnLeuGlnValLeuGlyLeuProHisAsnThr 180
 DB 481 GGGTTATGACCGCGGACGAGGCGCAAGAGTTTGAAMACCTACATCTCATACAAAG 540
 QY 181 PheThrValProTyrValIlePheAlaAsnLeuSerMetLeuAlaTyrLeuGlyArg 200
 DB 541 TACTGTGCTCCCTGCTGTCTGTCTTCACACTGGCGGACAGCGCCGAGGAGCGCC 600
 QY 201 IleArgAspThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgThrGln 220
 DB 601 ATCCGCGACACAGCGCCCTTAAAGCTGCTGCTCGAGGAGCTGAATGTTTTCGGGCG 660
 QY 221 CysGlyGlnLeuTyrAlaTyrAspTyrPheSerIlePheProLeuValTyrThrGlnVal 240
 DB 661 TGTGGAGTGTCTTACATATGATGATGATGATGATGATGATGATGATGATGATGATG 700
 QY 241 ThrValAlaValTyrSerPhePheLeuAlaCysLeuIleGlyArgIlePheLeuAsnPro 260
 DB 721 ACCATGCGACGTACAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
 QY 261 AsnLysAspTyrProGlyHisGlyMetLeuValValProValPheThrIleLeuGln 280
 DB 781 GCTCAGGGTTACAAAGACACAGCTAGACCTGTGTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 281 PheLeuPheTyrMetGlyTyrPheLeuValValGlnGlnLeuLeuAsnProPheGlyGln 300
 DB 841 TTTCTTCTTACAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 QY 301 AspAspAspAspPheGlyThrAsnTyrPheIleAspArgAsnLeuGlnValSerLeuLeu 320
 DB 901 GACGATGATGATCTTACAGCACTTCTGATGATGATGATGATGATGATGATGATGATG 960
 QY 321 SerValAspGlyMetHisGlnAsnLeuProPheMetGlnArgAspMetCysTyrTrpAsn 340
 DB 961 CAGGTGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 341 AlaAlaProGlnProProTyrThrAlaAlaSerAla---ArgSerArgArgHisSerPhe 359
 DB 1021 GCGGAGGCTGCGCGCCCATACACAGCGCTACTGCTTCCAGCGCGGACACCTTCCCT 1080
 QY 360 MetGlySerThrPheAsnIleSerLeuLysGlnValLeuTyrPheSerLysGln 379
 DB 1081 CAGGCTCTACCTTGTACATCAGCTGCGCCAAAGAACATGACAGTTCCAGCGGCTG 1140
 QY 380 GlnAlaAspThrAspLysLysGln-----Ser 388
 DB 1141 GCGTTGGATGAGCAGATGAGGAGCGCGCGGACACTCTCGAGCGCCCTTCCCGGCG 1200
 QY 389 GlyTyrSerSerThrIleGlyCysPheLeuGly-----LeuGlnProLys 403
 DB 1201 GCGCGGCGCATGTCGGGAGAGCGCCCTGCGCGCGCGCTGCTTCTTCTACCTCCCA 1260
 QY 404 AsnTyrHisLeuProLeuLysAspLeuLysThrLysLeuLeuGlySerLysAsnProLeu 423
 DB 1261 AAC-----AGTGTGCTGTGGAGCGCTACTGAGGCGCAGTGTCTCATACCGCGGT 1314
 QY 424 LeuGlnGlyGlnCysLysAspAlaAsnGlnLysAsnGlnLysAspValTrpLysPheLys 443
 DB 1315 CCCGAAAGCGCGCGCGCGAGTGCAC-----TGC 1344
 QY 444 GlyLeuAspPheLeuLysCysValProArgPheLysArgArgGlySerHisCysArgPro 463
 DB 1345 GGGGAGCCCGCTGTGACCC--CGGCTGTGCGGAGCGCGAGGCGCGCGCGCGCGCGG 1403
 QY 464 GlnAlaProSerSerHisPheProThrGlnGlnSerAlaProSerSerAspThrGlyAsp 483
 DB 1404 CGAACCGCTTACCTCATCCCTGCGGCGTG--TCGAGCCCTTACAGATGATGATGATG 1462


```

Db      882 GGC-----TTCAAGTCCGCTGCTGTGTGGAGCGCGGTTACCAAGCGGCCCA 932
Qy      464 GlnAlaProSerSerHis-----ProthGluGlnSerAlaPro----- 476
Db      933 CAGACACCCCTCGGCGACACCCCTATGCTTCCACCCGAAAGAGTCAGCGCCCTTAGGT 992
Qy      477 -----SerSerSerAspThrGlyAspGlyProSerThr 487
Db      993 CTTCGAGAGTCTCGGGCATAGATAGAGCTGCCAAGAACAGAGAGCTTCAGCTCGAGCT 1052
Qy      488 -----AspTyr 489
Db      1053 CCCAGCATCAGAGAGATTGAGTTGCTCCACAGAGAGCGCCGAGCCCTGCGGAGCC 1112
Qy      490 GlnGluIleCysHisMetIleGlySerHisValGluPheAsnLeu---AsnIleProGlu 508
Db      1113 CTGCAAGGAGTACACGTAGAGAGAGAGAGAGGATTAACTGCGCCGATCTGTGGAG 1172
Qy      509 SerProthGluHisLeuGlnGlnArgArgLeuAspGlnMetSerThrAsnIleGlnAla 528
Db      1173 GCGCCCG---GAAACATCTCAAGAACCGAATTGTGAGCA---CCGATGGGCAATACAGCT 1226
Qy      529 LeuMetIleGluHisAlaGluSerTyr-----ProTyrArgAspGluAla 543
Db      1227 ATACTCAAGACCCAGCGGATCCCTACTGGGCTTGGAAGAACAGGATGAGCA 1280

RESULT 11
AK000139
LOCUS      AK000139      1861 bp      mRNA      linear      PRI 22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20132 fls, clone COL06441.
ACCESSION AK000139
VERSION    AK000139.1 GI:7020032
KEYWORDS   oligo capping, fls (full insert sequence).
SOURCE     Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL06441.
ORGANISM   Homo sapiens
REFERENCE  1 (sites)
AUTHORS    Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,
           Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
TITLE      NEDO human cDNA sequencing project
JOURNAL    Unpublished
          2 (bases 1 to 1861)
          Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,
          Shibahara, T., Tanaka, T. and Nakamura, Y.
          Direct Submission
          Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
          University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
          Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
          Tel:81-3-5449-5286, Fax:81-3-5449-5416)
          NEDO human cDNA sequencing project supported by Ministry of
          International Trade and Industry of Japan; cDNA full insert
          sequencing; Research Association for Biotechnology; cDNA library
          construction; 5'- and 3'-end pass sequencing; Department of
          Virology and Human Genome Center, Institute of Medical Science,
          University of Tokyo (partly supported by Science and Technology
          Agency).
FEATURES
SOURCE     1. 1861
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /clone="COL06441"
           /clone_type="COL"
           /clone_lib="COL"
           /note="cloning vector pME18SFL3"
           65..1465
           /note="unnamed protein product"
           /codon_start=1
           /protein_id="BA090970.1"
           /db_xref="GI:7020033"
           /translation="MALSAAYRFVLTGEGKRFYKFLVYICDQYASLIPSVFLGYYT

```

```

LTVNRMMSSOICAPLDALMCVAVAGTVHGRDGRGLRYRTLRVAGLSAVLLRSVST
AVKRPRTIDHYVAGPMTREKRFENLNSSTNMTWPCWFSNLAQAREGRIRD
NSAKLLILELNVNFKGCMLEHYDMLSVLYTYOVTALYSFLACIGRQPLDPA
OGKRDLDLCYFTFTLDFEYAGNLKAEOLINPFGEDDDDFENFLIDNRQVSM
LAVDENTDGLAVLEKDLWDAAEAPAPYAAVFPOLROPSPFGSTFDITLKEKQFO
RUGLDGPMGAGARDLQRLILPAGAMGVAGVGLGRRLSPILBKNSCVSEASGASC
AVVPEGADECSGDPDLDPGLPPEAPAPPAPRPLTLIPGVPEFSLVTMPGPRGA
PMLSPIGEEENLA"
BASE COUNT      345 a      578 c      535 g      403 t
ORIGIN

Alignment Scores:
Pred. No.:      1,34e-103      Length:      1861
Score:          1147.50      Matches:      212
Percent Similarity: 76.99%      Conservative: 59
Best Local Similarity: 60.23%      Mismatches: 80
Query Match:      38.98%      Indels:      1
DB:              9      Gaps:      1

US-09-622-964-29 (1-551) x AK000139 (1-1861)
Qy      24 TTPArgGlySerIleTyrIleLeuLeuTyrGlyGluPheLeuValPheIlePheLeuTyr 43
Db      5 TGGCGTGGAGACATCTCAAACTCTGTGGAGAGAGCTGCTGTGCTTGGGTTCTAC 64
Qy      44 TyrSerIleArgGlyLeuTyrArgMetValLeuSerSerAspGlnGlnLeuLeuPheGlu 63
Db      65 ATGGCGCTGAGTGTGCTCCCTACCGCTTGTGCTGACCAAGAGAGAGAGAGAGAGAGAG 124
Qy      64 LysLeuAlaLeuTyrCysAspSerTyrIleGlnLeuLeuProIleSerPheValLeuGly 93
Db      125 AAGCTGTGATTATTTGTCAGCAGTANCCAGGCTATCCCTCTCTCTGCTGCTGCTG 184
Qy      84 PheTyrValIleLeuValIleSerArgTTPRPserGlnTyrGluAsnLeuProTyrPro 103
Db      185 TTTTATGTGAGCTGTGGTGGTGAACCCCTGTGTGAGCCAGTACCTATGATGCGCTGCC 244
Qy      104 AspArgLeuMetIleGlnValSerSerPheValIleGlyIleAspGlnGlnLeuTyrLeu 123
Db      245 GAGCGCCATATGTCGTGGTGGGCGGCGACCGTGCACGAGGAGAGAGAGAGAGAGAG 304
Qy      124 LeuArgArgThrLeuLeuArgTyrAlaIleLeuGlyGlnValLeuLeuLeuLeuSerIle 143
Db      305 TACCGCGACACTCATGCTGCTACGACGAGGCTCTGCGGCTCATCTGCTGCTGCTGCT 364
Qy      144 SerThrSerValTyrIleAspPheProThrIleHisIleValLeuAlaGlyPheMet 163
Db      365 AGCAGCGCGGTCTTCAAGCGCTTCCCAACATAGCCACGCTGTGGAGCGGTGGTTATG 424
Qy      164 ThrHisGlyGlnHisLeuGlnLeuGlnIleLeuGlyLeuProHisAsnThrPheThrVal 183
Db      425 ACCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
Qy      184 ProTyrValTyrPheAlaAsnLeuSerMetIleValTyrIleGlyIleArgIleAsp 203
Db      485 CCTTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
Qy      204 ThrValLeuLeuGlnSerLeuMetAsnGlnValCysThrLeuArgThrGlnCysGlyGln 223
Db      545 AACAGCGCCCTTAGCTGCTGCTGAGAGAGCTGAAATTTTGGGGGCAAAATGTGAAAG 604
Qy      224 LeuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnValAlaIleValAla 243
Db      605 CTCTTCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
Qy      244 ValTyrSerPhePheLeuAlaCysLeuIleGlyArgGlnPheLeuAsnProAsnIleAsp 263
Db      665 CTGTACAGCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724
Qy      264 TyrProGlyHisGluLeuValAlaProValPheThrIleLeuGlnIleLeuPhe 283
Db      725 TACAAGACACGACCTAGACCTGTGTGTGCGCATCTTCCATCCCTCTGAGTTCTTCTTC 784

```

QY 284 Tyrmctglttprpleuylsvalaglunleuleanpropheglygluasaspasp 303
DB 785 TACGCCGGCTGGCTCAGAGCTGAGCAGCTCATCAACCCCTCGAGAGACAGATAT 844
QY 304 AsphegluthtasripilleasparghansleuvalserleuleanserValasp 323
DB 845 GACCTTGAGACCACTTTCATGATGAAACTTCACAGCTGCATGCTGCGCAGTGAC 904
QY 324 GlymethisgluasleuProPromeCgluarqasmethyrttPaspagualalalPro 343
DB 905 GAGATGTATGATGACCTGGCTGTGAGAGAGACTTGTATGGATGCAAGCCGAGCT 964
QY 344 GlnProPrtYrttPrtalalaserAla---ArgSerArgArgHisSerPheMetglYser 362
DB 965 CGGCCCCATACACAGGCGCTACTGCTTCCAGCTGCGCAGCCCTTCCAGAGGCTC 1024
QY 363 ThPheasnlleSerleuYslysgluaspleuGl 374
DB 1025 ACCCTTGACATCAGCTGCGCCAAAGAACATGAC 1060

RESULT 12
LOCUS AF218817 2874 bp mRNA linear INV 09-FEB-2000
DEFINITION Drosophila melanogaster Dbest mRNA, complete cds.
ACCESSION AF218817
VERSION AF218817.1 GI:6942150
KEYWORDS
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2874)
AUTHORS Tavaani,B.C., Pappu,K. and Mardon,G.
JOURNAL dbest, a Drosophila Homolog of Human Bestrophin
TITLE unpublished
AUTHORS 2 (bases 1 to 2874)
JOURNAL Tavaani,B.C., Pappu,K. and Mardon,G.
TITLE Direct Submision
AUTHORS Submitted (23-DEC-1999) Department of Pathology and Program in
JOURNAL Developmental Biology, Baylor College of Medicine, One Baylor
Plaza, Houston, TX 77030, USA
FEATURES
source location/Qualifiers
1. 2874
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3"
/map="85F-86A"
411. 2576
/codon_start=1
/product="Dbest"
/protein_id="AAF32327.1"
/db_xref="GI:6942151"
/translation="MTTITGEVATCGRCGFCFLKLLRMGSTYKLVLDLAFLLTY
YAINMYRFGINPAKCTEPFAIYOCDSTRELPLSVLFYVSIWTRMMNOYTSIP
WPDIAFVSNNHGDGRMRMTIMRVCCLTMVLANVSPRYVKRFGNNLVE
AGLINDKETTITENKAFPRPSKHMPIYMAASITTRAKERIRDDPAVKITIEL
NKRQCGGLISYDTSVPLVYQVVTLYAYSLICCMCOQTDKRVAGTKTIEL
DLYPEVTTLQFFYMGWLKVAESLINFEEDDDFEVMMNDNIOVAYLIDENH
DHELLKDOYMDVEFNEPLPYTIAAEFRFRENHPEPSTAKIEYKNAKMSVSSVID
EMADDAAGIHFSAGNGKMRDSSPSLYSGTISRNTYASALKRFLSDSDSPGAT
PSODOPKPPFASASASISGAVYGSATSAKRPAGSRTITQVYEEVDEGATITSMAN
DPRPNWIDFAQTSAGSAGTSGLQPPPAISEVPDIPRSPTVNRASQSEPNLFFPG
VALLISTAPAGSPILLSNAATAPSSPVGESKSLYDQKGSRETVESMDLRSTD
LADDAVQDEDEDDFKLAEKEREKILRQOKNLARTISTAPGMEATVAPVAPV
NVAVQAOLOPVSADLLAGDQDFNSNTMSDAINGS"

BASE COUNT 777 a 700 c 721 g 676 t
ORIGIN

Alignment Scores:
Pred. No.: 7.16e-85 Length: 2874
Score: 961.00 Matches: 191
Percent Similarity: 68.874 Conservative: 70

Best Local Similarity: 50.408
Query Match: 32.648
DB: 3
Gaps: 6
US-09-622-964-29 (1-551) x AF218817 (1-2874)
QY 1 MetThIleThrThyThrAsnLysValAlaAsnAlaArg---LeuGlySerPheSer 19
DB 411 ATGACAAATTACGTACACAGAGTGAAGTGGCACTTGTCCGGCTTGGCTTTCTCAAA 470
QY 20 LeuLeuLeuCysTrpArgGlySerIleTyrlsLeuLeuTyrlsGlyluPheLeuAlaPhe 39
DB 471 TTGCTGCTCAGATGGGAGGAGCAATTACAAACGTTTGGCTGATCTTCTGGCTTC 530
QY 40 IlePheLeuTyrlsSerIleArgGlyLeuTyrlsArgMetValLeuSerSeraspIndln 59
DB 531 TTGACCAATTATATGCGATCAACATGGTGTATGCTTGGCTTACCCCGCAGCAAAA 590
QY 60 LeuLeuPheGlyLysLeuAlaLeuTyrlsCysaspSerTyrlleGlnLeuIleProIleSer 79
DB 591 GAAACCTTGGAGGCCATTGTTCAGTACTGTATGTTACAGAACTATACCCCTGTCC 650
QY 80 PheValIleuGlyPheTyrlsValThrLeuValAlaSerArgTrpTrpSerGlnTyrlsAsn 99
DB 651 TTCGTGCTTGGTTCTATGATCATCTGTGATGACCCGTTGGTGAATACATACACCTCC 710
QY 100 LeuProTrpProAspArgLeuMetIleGlnValSerSerPheValIleGlyLysAspGlu 119
DB 711 ATTCCTCGCCAGATCCCATCCAGCTCCGCTTGTTCAGTCAATGTCATGCGAGAGATG 770
QY 120 GluGlyArgLeuLeuArgArgThrLeuIleArgTyrlsAlaIleuGlyGlnValIle 139
DB 771 CGAGAGCAGCATGAGAGCCGCAACATATGATGTGCTTGGCTGACAGATGCTC 830
QY 140 LeuArgSerIleSerThrSerValTyrlsArgPheProThrLeuThsIleValIleu 159
DB 831 CTGGCGATGTTTCGCCAGAGGTGAAGAACGTTTCCCGCCCTAATATCTGCTGGA 890
QY 160 AluGlyPheMetThrThsIleGlyLysGlnLeuGlnIleuLeu-----GlyLeuPro 177
DB 891 GCGGCTGTGTAATGACAAATGAAGAACCATCATGAGAACATGAAGAGCCCTTCCC 950
QY 178 His---AsnThrPheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTyrl 196
DB 951 AGACCTTGCAGACGACTGGCTGCCCATGTTGGCGCCAGATTTTATACAGGGCCAGA 1010
QY 197 LeuGlyArgIleArgAspThrValIleLeuGlnSerLeuMetAsnIleValCysThr 216
DB 1011 AAGGAGGTGCGATTCGTATGATATTTGCTGTGAAGACCATCATGATGAGCTAATAG 1070
QY 217 LeuArgThrGlnCysGlyGlnLeuTyrlsTyrlsAspTrpIleSerIleProIleValTyrl 236
DB 1071 TTTCGTGGTCAAGTGGAGTCTCTCATACAGTACGATACCATATAGTGTGCTGTGAC 1130
QY 237 ThrGlnValValThValAlaValTyrlsSerPhePheLeuAlaCysLeuIleGlyArgGln 256
DB 1131 ACCCAAGGTGACACCTGCGGTGTATTCGATCTTACCTCTGATGCTGATGATACAA 1190
QY 257 PheLeuAsn-----ProAsnLysAspTyrlProGlyHisGlnMetAsnLeuVal 272
DB 1191 TGGACCGATGCGAAGTGGTGGCCAAATACCAATACCTG---AACAAAGTGTGATTAAC 1247
QY 273 ValProValPheThrIleLeuGlnPheLeuPheTyrlsGlyTrpLeuValAlaGlu 292
DB 1248 TTTCCTGATTTACACAGCTCAGTCTTCTTCAATAGGTTGGCTCAAGTGGCCGAC 1307
QY 293 GlnLeuIleAsnPropheGlygluasaspaspPhegluthtasripilleasp 312
DB 1308 TCGCTGTAATTCATTTGGGAGAGCATGATATTTTGGAGTCAATGATGCTGAT 1367
QY 313 ArgAsnLeuGlnValSerLeuLeuSerValaspGlyMetHisGlnAsnLeuProProme 332
DB 1368 CGCAATCTTCAAGTGTCTCTATCTGATGCTGAGACAGATGACCATGACATCCGAGACTG 1427

QY 333 GUATGAPMETYTRTPASNGUUAAlaAPro---GlnProProTyThrAlaAlaSer 351
DB 1428 TTAAGCATCTACTGCGAGCGAGGTGTTCCTCCCAAGCGAGCGCTTACACATAGCTGCC 1487
QY 352 AlAaGSeArArHlSerPheMetGlySerThrPheAnIleSerLeuLysLys 370
DB 1488 GAACGATTCCGGAGATCATCCAGACCGCTCCACTGCCAAGATCGAGTCCCAAG 1544
RESULT 13
AY061546. 5610 bp mRNA linear INV 08-NOV-2001
LOCUS Drosophila melanogaster LP05915 full length cDNA.
DEFINITION AY061546
ACCESSION AY061546
VERSION AY061546.1 GI:16769749
KEYWORDS FLI-CDNA.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5610)
Stapleton, M., Brokslein, P., Hong, L., Agdayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Fattan, D., Filse, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S. E., Rubin, G. M. and Celisner, S.
Direct Submission
Submitted (30-OCT-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
FEATURES
Source location/Qualifiers
1..5610
/strain="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
1..4986
/note="bias alignment with AE003686"
3142..5307
/note="Longest ORF"
/codon_start=1
/product="LP05915p"
/protein_id="AAL29094.1"
/db_xref="GI:16769750"
/translation="MTITGEGVATCRGFCGLKLLRMGSIYKLVMLDLAFETY
VAINVVRGRLNPAOKETFEALIVQCDSEILPLSLVGLAYSVIATRMNYSIP
WDPPIAVSVSNVHODERGRMRRTIRVAVCLCLTAVLANVSRVKKRRGNINYE
NKRFGQCLSLIDYTLISVPLVYOVVLAVSYSLCCMGCGOIRDPVAKTIDEL
DLIFPVTTLQEFYMGMLKVAESLINPFGDDDDPPVNMVNRNLDVSLVIDENH
DHELIDQYWDVEFNPMLPYTLIAERFRNRPSTAKLEVRNAMPSTMSVSD
EMADASGIIHFSAGNKRILDDSSPSLYSGTISRAKRVASAKRLSRDSRSGSAT
PSODQPKFPASASASLSCAVYGSATSAKPAQSALITQOVLIEVLEADQATISMAN
DPRNVNIDFAQTSAGTSGPIQPPAHSEPDIDPSRPSYNRAQSOREPNLPPGCG

Alignment Scores:
Pred. No.: 1,81e-84 Length: 5610
Score: 961.00 Matches: 191
Percent Similarity: 68.87% Conservative: 70
Best Local Similarity: 50.40% Mismatches: 108
Query Match: 32.64% Indels: 10
DB: 3 Gaps: 6
US-09-622-964-29 (1-551) x AY061546 (1-5610)
QY 1 MethrilleThrTyThrAnLysValAlaAlaAlaArg---LeuGlySerPheSerSer 19
DB 3142 ATGACATTTACGTACACAGGTGAGAGTGGCCACTTGGCGGCTTGGCTTTCACAA 3201
QY 20 LeuLeuLeuGlySerTyArgGlySerIleTyArgLeuLeuTyArgLuphLeuValPhe 39
DB 3202 TTGCTGCTCAGATGGCGAGACCAATTTACAAACTGGTTGGCTTACATCTTGGCTTC 3261
QY 40 IllePheLeuTyTySerIleLeuArgGlyLeuTyArgMetValLeuSerSeraspGlnGln 59
DB 3262 TTGACCATTTACTATCGATCAACAGTGGTATCGCTTGGCTTCAACCCCGCACAA 3321
QY 60 LeuLeuPheGlyLeuLeuAlaLeuTyCysAspSerTyIleGlnLeuLeuProIleSer 79
DB 3322 GAACCTTGGAGCCATTTGTCAGTCTGATTTACAGAGAACTATACCTGTC 3381
QY 80 PheValLeuGlyPheTyValPheLeuValAlaSerArgTPTPTSerGlnTyGlnAsn 99
DB 3382 TTGCGCTGCTGTTTCTATGATTCATGTGATGACCCCTGTGGTGAATCAGTACCTCC 3441
QY 100 LeuProTTPProAspArgLeuMetIleGlnAlaSerSerPheValGlyGlyAspGlu 119
DB 3442 ATTCCTTGGCGAGACCCATCGCGCTGTTGCTGATGTCATGTCATGGCCAGATGAG 3501
QY 120 GlnGlyArgLeuLeuArgTyThrLeuLeuArgTyArgAlaIleLeuGlyGlnValLeu 139
DB 3502 CGAGAGCCATGATGAGCCGACCAATGATGATGTGCTTGGCTGAGAGATGCTC 3561
QY 140 LeuArgSerIleSerThrSerValTyArgArgPheProThrLeuHSHSLeValLeu 159
DB 3562 CTGGCGAATGTTCGCGAGGAGTGAAGACGTTTCCCGGCTTAATATATCTGGTGA 3621
QY 160 AlaGlyPheMetThrHISGlyGlnHISLysGlnLeuGlnLysLeu-----GlyLeuPro 177
DB 3622 GCGGCTGCTGAATGATGCAATGAAGAACCATCATGACACCATGAACAGGCTTCCC 3681
QY 178 HIs---AsnThrPheThrValProTyValTTPPheAlaAsnLeuSerMetLysAlaTy 196
DB 3682 AGACCTTGAAGACGACGCGGCGCCATGCTTGGCTGCGAGTATTTAAACGAGGCGA 3741
QY 197 LeuGlyIleArgIleArgAspThrValLeuLeuGlnSerLeuMetAsnGluValCysThr 216
DB 3742 AAGGAGGTGCGCTTGTGATGATTTTGTCTGGAAGACCATCATGATGATGATGA 3801
QY 217 LeuArgTyGlnCysGlyGlnLeuTyArgAlaTyArgAspTTPPheIleProLeuValTy 236
DB 3802 TTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3861
QY 237 ThrGlnValAlaValValAlaValTySerPhePheLeuAlaCysLeuIleGlyArgGln 256
DB 3862 ACCCAAGTGGTACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3921
QY 257 PheLeuAsn-----ProAsnLysAspTyTyProGlyHISGlnMetAspLeuVal 272
DB 3922 TGGACCGATGCAAGGTGGTGGCAATATCAACATCACTG---AACAGGTGATCTATAC 3978
QY 273 ValProValPheThrIleLeuGlnPheLeuPheTyMetGlyTyThrLeuLysValAlaGln 292

```

Db      3979  TTTCTGATTTTAAAGCTGCGACATCTTCTTACATGGTGGCCCAAGGCGCGAG 4038
Qy      293  GlnleuileanProPieglYgluAspAspAspPheglurThrAntPileileasp 312
Db      4039  TCGCTGATTAATCCATTTGGCGAAGCATGATGATTTTGAGTCAACAGATGGAT 4098
Qy      313  ArgasnleuGlnValSerleuLeuSerValaspGlyMetHisGlnasnleuProPhe 332
Db      4099  CGCATCTTCAAGTGTCTATCTGATCTGCAGACATGCACCATGACCATCGAGCTG 4158
Qy      333  GluArgaspMetIleTyrPasnGluAlaAlaPro---GlnProPtyrThrAlaLase 351
Db      4159  TTAAGAGATCAGACTGCGACGAGCTGCTCCCAAGAGAGCTGCCCTACACATAGCT 4218
Qy      352  AlaArgSerArgIleSerPheMetGlySerThrPheAsnIleSerleuLys 370
Db      4219  GAACGATTCGGGAGAAATCATCAGAGCCGTCCTGCGCAAGATGAGTCCCAAG 4275

RESULT 14
AC019521
LOCUS      32022 bp      DNA      linear      HTG 03-JAN-2000
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, In ordered
ACCESSION  AC019521
VERSION     AC019521.1  GI:6665376
KEYWORDS   HTG: HTGS_PHASE2.
SOURCE      Drosophila melanogaster.
ORGANISM   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
            Adams, M. and Venter, J.C.
            Direct Submission
            Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
            This sequence was identified as CDM:1021063 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

FEATURES
    source             1..32022
                        /organism="Drosophila melanogaster"
                        /db_xref="taxon:7227"
                        /locus_tag="CG1021063"

BASE COUNT  9830 a      6651 c      6507 g      9034 t

Alignment Scores:
Pred. No.:      2.39e-70      Length:      32022
Score:          829.00      Matches:      194
Percent Similarity: 48.46%      Conservative: 73
Best Local Similarity: 35.21%      Mismatches: 120
Query Match:      28.16%      Indels:      165
                        Gaps:      11

US-09-622-964-29 (1-551) x AC019521 (1-32022)
Qy      21  LeuLeuGlyTyrPargIleSerIleTyrIleLeuLeuTyrGlyGlnPheLeuValPhe 40
Db      5825  CTTTTCAGATGGCGAGGAACATTTACAAACATGGTTGGCTAGATCTTCTGCGCTT 5884
Qy      41  PheLeuTyrTyrSerIleArgGlyLeuTyrArgMetValLeuSerSerSpInln--- 59
Db      5885  ACCATTACATGCGATCAACATGATGCTGCTTGGCTTAAACCCGCAAAAAAGAG 5944
Qy      59  -----
Db      5945  TGAATATAAAGCTCCTAGAAAGATCATGAGTGTGTGCTCATTAACCAT 6004
Qy      59  -----

Db      6005  TTTATGATTAATGAATATATATATTTTGAATTTATCATATGTTCCATTAATTCGAA 6064
Qy      59  -----
Db      6065  AGTCTGATGAGGCTCTAATCAATATCAATATGATCAATTTAAATACAGACAGTA 6124
Qy      59  -----
Db      6125  GAAAGTTTACTCAAAACATCTGAAATATAGCTTTCAGCACTTCACAAATGCTTTTAA 6184
Qy      60  -----
Db      6185  AAAATTAGCTGATCTACAAATATACATATATATATGATGATGATCTTCTCCACGA 6244
Qy      62  -----
Db      6245  TCTTGTCTCTTTTAAAGCGCTTACTATAATGCAACTTCTTACAGACCTTGAGGCC 6304
Qy      65  LeuAlaLeuTyrCysAspSerTyrIleGlnleuIleProIleSerPheValleuGlyPhe 84
Db      6305  ATTGTTCAGTACTGATGATAGTACAGACAGACATCCCTGCTCTGCTGCTGTTTC 6364
Qy      85  TyrValThrPheValIleSerArgTyrTrpSerGlnTyrGluAsnleuProThrAsp 104
Db      6365  TATGTATGATGTGTGATGACCCGTTGGTGAATCATGATACCTTCATTCCTGGCCAGAT 6424
Qy      105  ArgIleMetIleGlnIleSerSerPheValGlnGlyIleValleuLeuLeuArgSerIleSer 144
Db      6425  CCCATCGCGGCTTTGTCAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 6484
Qy      125  ArgArgThrIleLeuLeuTyrValIleLeuGlyIleValleuLeuLeuArgSerIleSer 144
Db      6485  AGCGAACAATTAAGCCATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6544
Qy      145  ThrSerValTyrIleAspArgPheProThrIleHisIleValleuValleuIlePheMetThr 164
Db      6545  CCGAGGAGTGAAGAGCGTTTCCCGGCGCTTAATATCTGTGAAGCGGCTGCTGAAT 6604
Qy      165  HisGlyGlnHisIleGlnleuGlnIleLeu-----GlyLeuProHis---AsnThrPhe 181
Db      6605  GAAATGAAGAAAGACATCATGACAGACATGACGAAACAGCGCTTCCAGACCTTGAAGAC 6664
Qy      182  TrpValProTyrValTyrPheAlaIleAsnleuSerMetIleValTyrLeuGlyValTyrGly 201
Db      6665  TGGCTGCCATGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6724
Qy      202  ArgAspThrValleuLeuGlnSerleuMetAsnGluValIleCysThrIleuArgThrGlnCys 221
Db      6725  CGTATGATTTTGTCTGTAAGACATCATGATGATGATGATGATGATGATGATGATGATG 6784
Qy      222  GlyIleLeuTyrAlaTyrAspTrpIleSerIleProleuValTyrTrpGlnValIleThr 241
Db      6785  GAGCTCTCTCAGACAGTACAGATACCAATGATGATGATGATGATGATGATGATGATGAT 6844
Qy      242  ValAlaValTyrSerPhePheLeuAlaCysIleuIleGlyIleArgGlnPheLeuAsn----- 259
Db      6845  CTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 260
Qy      260  -----ProAsnIleAspTyrProGlnHisGlnMetAspLeuValProValIlePheThr 277
Db      6905  GTGGTGGCAATACACATACCTG---MCMAAGTGGATCTAATACCTTCTGATTTTAA 6961
Qy      278  IleLeuGlnIlePheLeuPheTyrMetGlyTyrPheLeuValAlaGlnleuIleAsnPro 297
Db      6962  ACCCTGCAAGTCTTCTTCTTACATGATGATGATGATGATGATGATGATGATGATGAT 7021
Qy      298  PheGlyIleAspAspAspPheGlu-----
Db      7022  TTTGGCGAAGACATATATATTTTGAAGTACCATTAACACATTTGTAGTCTGCGA 7081
Qy      307  -----ThrAsnTrpIleIleAspArgAsnleuGlnValSer 318
            |||||-----|||||

```


Search completed: July 27, 2003, 15:02:22
Job time : 4312.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:

July 27, 2003, 11:10:16 ; Search time 1737.51 Seconds
(without alignments)
4054.679 Million cell updates/sec

Title:

Perfect score:

Sequence:

Scoring table:

US-09-622-964-5
2350
1 MTTTTSQVANRLRSGFSRL.....QLLPSSVEPLMAPSTSA 435
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched:

16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:

32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-CGN2.1/USPTO_SPOOL/US09622964/runat.22072003_101119_26026/app_query.fasta_1.2069
-DB=EST -QFMT=faststep -SUFFIX=first -MINMATCH=0.1 -LOOPC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTWMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09622964.BCGN.1.1.4186.grunat.22072003_101119_26026 -NCPU=6 -ICPU=3
-NO_MAP -LARGESOURCE -NEB_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1503.5	64.0	2453	11	BC0152220 Homo sapiens
2	1318.5	56.1	963	11	BQ0798800 AGENCOURT
3	1218	51.8	1699	11	AK006549 Mus musculus
4	957	40.7	735	13	B1756228 B1756228
5	940	40.0	666	13	B1480798 B1480798
6	799	34.0	522	14	BM707649 BM707649
7	764	32.5	566	14	BM718338 BM718338
8	750	31.9	793	14	BM718338 BM718338
9	746	31.7	525	14	BQ345562 BQ345562
10	745	31.7	526	14	BM685501 BM685501
11	699	29.7	613	9	AL654606 AL654606
12	698	29.7	508	10	BE236901 BE236901
13	674	28.7	1067	13	BM562042 BM562042
14	673.5	28.7	722	9	AL627604 AL627604
15	662	28.2	628	13	BQ069061 BQ069061
16	649	27.6	732	13	BQ65806 BQ65806
17	614	26.1	657	14	BQ397157 BQ397157
18	604.5	25.7	603	9	AA205892 AA205892
19	588	25.0	654	10	BE663909 BE663909
20	578	24.6	577	10	BE663909 BE663909
21	574	24.4	531	13	B1343182 B1343182
22	574	24.4	514	13	BM685396 BM685396
23	572	24.3	966	12	BE582222 BE582222
24	570.5	24.3	966	12	BE582222 BE582222
25	567	24.1	1111	17	CNS0402M CNS0402M
26	550	23.4	635	14	BQ396082 BQ396082
27	549	23.4	649	10	BE189780 BE189780
28	535	22.8	605	14	BQ391968 BQ391968
29	534.5	22.7	730	13	BQ963329 BQ963329
30	532	22.6	1123	14	BQ76342 BQ76342
31	498	21.2	308	12	BG892937 BG892937
32	471	20.0	797	10	AV398472 AV398472
33	464	19.7	614	9	AL782195 AL782195
34	457.5	19.5	566	13	B1443895 B1443895
35	453	19.3	470	9	AA573517 AA573517
36	448	18.5	305	9	AA318352 AA318352
37	434.5	18.5	509	13	BQ030264 BQ030264
38	427	18.2	685	13	BQ126118 BQ126118
39	426	18.1	972	17	CNS042X4 CNS042X4
40	421.5	17.9	672	13	BM576476 BM576476
41	420	17.9	930	14	BQ436824 BQ436824
42	419	17.8	1033	17	CNS040M2 CNS040M2
43	413	17.6	440	12	BE084076 BE084076
44	409	17.4	671	9	AL598355 AL598355
45	408	17.4	613	9	AU208590 AU208590

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
BC015220	BC015220	Homo sapiens, clone IMAGE:3877806, mRNA.	BC015220	GI:21955361	HTC.	Homo sapiens.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 2453)	Strausberg, R.	Submitted (01-OCT-2001)	National Institutes of Health, Mammalian
BC015220	BC015220	Homo sapiens, clone IMAGE:3877806, mRNA.	BC015220	GI:21955361	HTC.	Homo sapiens.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 2453)	Strausberg, R.	Submitted (01-OCT-2001)	National Institutes of Health, Mammalian

DNA Sequencing by: Agencourt Bioscience Corporation
 found through the MGC clone distribution information can be
 http://image.llnl.gov
 Plate: LLM13563 row: 1 column: 08
 High quality sequence stop: 623.
 Location/Qualifiers

FEATURES

1. 963
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="6180559"
 /clone_id="Lupski_dorsal_root_ganglion"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site:1:
 NotI; Site:2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCAGCGCTCCG-3' and
 5'-GACTAGTTCAGATCGGAGCGCCGCTT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 Technology and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 205 a 289 c 265 g 202 t 2 others
 ORIGIN

US-09-622-964-5 (1-435) x B0879880 (1-963)

Alignment Scores:
 Pred. No.: 1 49e-105 Length: 963
 Score: 1318.50 Matches: 251
 Percent Similarity: 77.98% Conservative: 4
 Best Local Similarity: 76.76% Mismatches: 2
 Query Match: 56.11% Indels: 70
 DB: 14 Gaps: 2

Db 493 GACCTGCTTGGCCCTTCACGTTCTCTCTATGTTGGCTGGAAG 552
 Qy 289 sValGlyLeuSerArgAlaLeuLeuGlyTrpArgHisGlyGlnGlyGlnGln 309
 Db 553 G----- 553
 Qy 309 nLeuLeuGluTrpArgMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 329
 Db 553 ----- 553
 Qy 329 nAlaTrpTrpArgTrpProValIleProAlaThrArgGlnAlaGlnGlnGln 349
 Db 553 ----- 553
 Qy 349 uGluProGlyArgArgGlyLeuTrpTrpGlnSerSerSerSerThrProLeuGluArg 369
 Db 554 ----- 554
 Qy 369 tMetMetIleLeuArgProThrGlyLeuSerThrGlyIleGlyArgCysProCysTrp 389
 Db 589 GATGATGATTTTGGACCACTGATTTGACAGGAAATTTGCAGGTGTCCTGTTGGCT 648
 Qy 389 uTrpMetArgCysThrArgThrCysLeuGlyTrpSerArgTrpCysThrGlyIleSer 409
 Db 649 GTGATGATGATGACACGACGACCTGCTCGGATGACCGGACATGACTGAAATAGCCC 708
 Qy 409 oSerHisSerProProThrGlnLeuLeuProProSerSerValGluProLeuTrpAl 429
 Db 709 GAAGCACAGGCCCCCTACACAGCTGTTCCGCCAGAGTTCGTCGAGCTCTTATGAGGC 768
 Qy 429 aProProSerThrSerAla 435
 Db 769 TCCACCTTCAACATCAGCC 787

RESULT 3
 AK006549
 LOCUS
 DEFINITION
 Mus musculus adult male testis cDNA, RIKEN full-length enriched
 library, clone:1700030H21:vitaliiform macular dystrophy 2 homolog
 (human), full insert sequence.
 ACCESSION
 AK006549
 VERSION
 AK006549.1 GI:12839710
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (strain: C57BL/6J) adult male testis cDNA to mRNA,
 clone:1700030H21.
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 REFERENCE
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL
 MEDLINE
 PUBMED
 99279253
 10349636
 AUTHORS
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 11042159
 3
 REFERENCES
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishi, Y., Nakamura, S., Hazama, N., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencing

Db 699 ----- 699
 QY 316 nCysGlnGluArgValSerArgValGluSerSerGlnAlaTrpTrpArgThrProva 336
 Db 699 ----- 699
 QY 336 lIleProAlaThrArgGluAlaGluAlaGlyGluSerLeuGluProGlyArgArgGly 336
 Db 699 ----- 699
 QY 336 uTrpTrpGlnSerSerSerSerThrProLeuGluArgMetMetIleLeuArgProTh 376
 Db 700 ----TGGCAGAACGCTCATCAACCCCTCGGGAGACAGATGATTTTGAAGCTAAC 755
 QY 376 rGlyLeuSerThrGlyIleCysArgCysProCysTrpLeuTrpMetArgCysThrArgTh 396
 Db 756 TGGATCATTTGACAAACCTGCAGAGTGCTCCCTGTTCCTGGATGGATGACACAGAAC 815
 QY 396 rCysLeuGlyTrpSerArgThrCysThrGlyIleSerProSerHisSerProProThrG1 416
 Db 816 TTGGCTCCATGAGAGCTGACATGCTGTGAGAACGACGACGCGCTCAGCCGCGCTACAC 875
 QY 416 nLeuLeuProProSerSerValGluProProLeuTrpAlaProProSerThrSerAla 435
 Db 876 GCTGCTTCTGCGAGTCTCGCGGCGATCTTTCATGGGCTCCACCTCAACATCAGCC 933
 RESULT 4
 LOCUS B1756228 735 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603024265F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194649 5',
 mRNA sequence.
 ACCESSION B1756228
 VERSION B1756228.1 GI:15747806
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 735)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINt at:
 http://image.llnl.gov
 Plate: LLAM1487 row: a column: 18
 High quality sequence stop: 577.
 Location/Qualifiers
 1. 735
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5194649"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1: NotI;
 Site: 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."
 135 a 229 c 214 g 157 t

Alignment Scores:
 Pred. No.: 6,23e-74 Length: 735
 Score: 957.00 Matches: 199
 Percent Similarity: 90.71% Conservative: 7
 Best Local Similarity: 88.05% Mismatches: 6
 Query Match: 40.72% Indels: 14
 DB: 13 Gaps: 2

US-09-622-964-5 (1-435) x B1756228 (1-735)

QY 52 LeuAlaLeuThrGlnGlnGlnGlnLeuMetPheGluLysLeuThrLeuTrpCysAspSer 71
 Db 60 CTGGCCCTCAGAGAAACAAACACAGTATGTTGAGAACTGACTCTGATTCGACAGC 119
 QY 72 TyrIleGlnLeuIleProIleSerPheValLeuGlyPheTrpValThrLeuValThr 91
 Db 120 TACATCCAGCTCATCCCATTCCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 179
 QY 92 ArgTrpTrpAsnGlnTrpGluAsnLeuProTrpProAspArgLeuMetSerLeuValSer 111
 Db 180 CGCTGGTGAACACAGTACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239
 QY 112 GlyPheValGluGlyLysAspGlnGlnArgLeuLeuArgArgThrLeuThrArgTrp 131
 Db 240 GGCTTGTGTGCAAGCAAGAGACAGCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
 QY 132 AlaAsnLeuGlyAsnValLeuIleLeuArgSerValSerThrAlaValTrpLysArgPhe 151
 Db 300 GCCAACCTGGCAGAGCTGCTCATCTGCGCAGCTGACACGCGAGTCTACAAAGCGCTTC 359
 QY 152 ProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeu 171
 Db 360 CCCAGCGCCACACACTGCTGTCAGACAGCTTATGATCCGCGCAACACACAGCTTG 419
 QY 172 GlnLysLeuSerLeuProHisAsnMetThrTrpValProTrpValTrpPheAlaAsnLeu 191
 Db 420 GAGAAACAGAGCTTCCACACAACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 479
 QY 192 SerMetLysAlaTrpLeuGlyLysArgIleArgAspProIleLeuGlnSerLeuLeu 211
 Db 480 TCAATGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
 QY 212 AsnGlnMetAsnThrLeuArgThrGlnCysGlyHisLeuTrpAla-TyrAspTrpIle 231
 Db 540 AACAGATGAACACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
 QY 231 erIleProLeuValTrpThrGlnValVal-----ThrValAlaValTrp 245
 Db 600 GTATCCCT-----ACTGGGTATTATCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
 QY 246 SerPhePheLeuThrCys-LeuValGlyArg-GlnPheLeuAsn-ProAlaLysAlaTrp 264
 Db 651 GGTTCCTTCCTGACTGTCTTGTGTGGGGGGGCTTGTGAACCCCGGCGGCGCTTC 710
 QY 265 ProGly 266
 Db 711 CCTGGG 716
 RESULT 5
 LOCUS B1480798 666 bp mRNA linear EST 28-FEB-2002
 DEFINITION H2RE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
 sequence.
 ACCESSION B1480798
 VERSION B1480798.1 GI:18998607
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Buraczynska,M., Mears,A.J., Zarepari,S., Farjo,R., Filipova,E.,

QY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 DB 60 ATGACCATCACTTACACAAAGCCAGTGGCTAATCCGCGCTTAGCTCTTCCCGCTG 119
 QY 21 LeuLeuGlySerPheGlySerIleTyrIleLeuLeuTyrGlyGlnPheLeuIlePheLeu 40
 DB 120 CTGGTGTGCTGGCGGCGGACAGCTACACAGCTCTATATGCGAGTTCCTAATCTCCG 179
 QY 41 LeuGlyTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGlnGlnGlnLeu 60
 DB 180 CTCTGCTACTACATCAATCCGCTTATTTATAGCTGCGCTTCAGGAAACACACAGCTG 239
 QY 61 MetPheGlnLeuLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
 DB 240 ATCTTGAGAAACTGACTCTGTATTCGACAGCTACATCAGCTCATCCCATTTCTTC 299
 QY 81 ValLeuGlyPheTyrValThrLeuValValThrArgTyrPheAsnGlnTyrGlnAsnLeu 100
 DB 300 GTGCTGGCTTACGTAGCGCTGCTGCTGACCCGCTGAGAACAGTACAGAACCTG 359
 QY 101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyAspGlnGln 120
 DB 360 CCGTGGCGCGGACCGCTCATGTAGCCCTGTGTGGGCTTCGTGAAGGACAGAGAGCA 419
 QY 121 GlyArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeu 140
 DB 420 GCCCGGCTGCTGGCGGACGCTCATCCGCTACGCCAACCTGGGCAACGTGCTCATCTG 479
 QY 141 ArgSerValSerThrAlaValTyrIleAspArgPheProSerAla 154
 DB 480 CCGACGCTACGACCGCGAGTCTACAAAGCGCTTCCCGACGCGC 521
 RESULT 7
 LOCUS BM718338 566 bp mRNA linear EST 01-MAR-2002
 DEFINITION UI-E-EO1-a1a-j-23-0-UI r1 UI-E-EO1 Homo sapiens cDNA clone
 ACCESSION BM718338
 VERSION BM718338.1 GI:19036537
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 566)
 TITLE Bernaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@iuiiuee.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 FEATURES
 source
 1.566
 location/qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-EO1-a1a-j-23-0-UI"
 /clone_1id="UI-E-EO1"
 /tissue_type="fetal eye"
 /dev_stage="fetal"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EO1 is a normalized cDNA library containing the
 following tissue(s): fetal eye. The library was
 constructed according to Bernaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 (dT)18 tail. The sequence tag for this library is
 CGCGTATAC. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI)."
 BASE COUNT 107 a 182 c 145 g 132 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,97e-57 Length: 566
 Score: 764.00 Matches: 158
 Percent Similarity: 61.008 Conservative: 0
 Best Local Similarity: 61.008 Mismatches: 1
 Query Match: 32,518 Indels: 100
 DB: 14 Gaps: 1
 US-09-622-964-5 (1-435) x BM718338 (1-566)
 QY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 DB 88 ATGACCATCACTTACACAAAGCCAGTGGCTAATCCGCGCTTAGCTCTTCCCGCTG 147
 QY 21 LeuLeuGlySerPheGlySerIleTyrIleLeuLeuTyrGlyGlnPheLeuIlePheLeu 40
 DB 148 CTGGTGTGCTGGCGGCGGACAGCTACACAGCTCTATATGCGAGTTCCTAATCTCCG 207
 QY 41 LeuGlyTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGlnGlnGlnLeu 60
 DB 208 CTCTGCTACTACATCAATCCGCTTATTTATAGCTGCGCTTCAGGAAACACACAGCTG 267
 QY 61 MetPheGlnLeuLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
 DB 268 ATGTTGAGAAACTGACTCTGTATTCGACAGCTACATCAGCTCATCCCATTTCTTC 327
 QY 81 ValLeuGlyPheTyrValThrLeuValValThrArgTyrPheAsnGlnTyrGlnAsnLeu 100
 DB 328 GTGCTGGCTTACGTAGCGCTGCTGCTGACCCGCTGAGAACAGTACAGAACCTG 387
 QY 101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyAspGlnGln 120
 DB 388 CCGTGGCGGACCGCTCATGTAGCCCTGTGTGGGCTTCGTGAAGGACAGAGAGCA 447
 QY 121 GlyArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeu 140
 DB 448 GCCCGGCTGCTGGCGGACGCTCATCCGCTACGCCAAC----- 486
 QY 141 ArgSerValSerThrAlaValTyrIleAspArgPheProSerAlaGlnIleValGlnAla 160
 DB 486 ----- 486
 QY 161 GlyPheMetThrProAlaGlnIleGlnLeuGlnIleGlnLeuSerLeuProIleAsnMet 180
 DB 486 ----- 486
 QY 181 PheTyrValProTyrValThrPheAlaAsnLeuSerMetCysAlaTyrLeuGlyArg 200
 DB 486 ----- 486
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuAlaGlnGln 220
 DB 486 ----- 486

QY 221 CysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValVal 240
 Db 487 -----CTGGATATATACAGAGGTGTG 507

QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsn 259
 Db 508 ACTGTGGCGGTGTACAGCTTCTTCTGTACTGTCTGTGTGGCGGCGCATCTTGTGAC 564

RESULT 8
 BG961794 793 bp mRNA linear EST 12-JUN-2001
 LOCUS 602826439f1 NCI_CGAP_C024 Mus musculus cDNA clone IMAGE:4981425 5',
 DEFINITION mRNA sequence.
 ACCESSION BG961794
 VERSION BG961794.1 GI:14349431
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 793)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloned by: Incyte Genomics, Inc.
 Cloned distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL0982 row: m column: 10
 High quality sequence stop: 765.
 Location/Qualifiers
 1..793
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_image="4981425"
 /clone_lib="NCI_CGAP_C024"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: colon; Vector: pCMV-Sport6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 1.6 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 170 a 221 c 213 g 189 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.13e-55 Length: 793
 Score: 750.00 Matches: 138
 Percent Similarity: 58.72% Conservative: 34
 Best Local Similarity: 48.32% Mismatches: 66
 Query Match: 31.91% Indels: 70
 Db: 13 Gaps: 2

US-09-622-964-5 (1-435) x BG961794 (1-793)

QY 77 ProIleSerPheValLeuGlyPheTyrValThrLeuValValThrArgTyrTrpAsnGln 96
 Db 13 CCGGCTCTTGTCTACTTGGCTTCTTCTGACTGTGTGTGCATGCGCTGGTGGAAACAG 72

QY 97 TyrGlnAsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGlnGly 116
 Db 73 TACCTATGCATCTCTGCGGAGCGACATGTCATAGTGGCTGGCAGCCGTGCATGGG 132

QY 117 LysAspGlnGlnGlyArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsn 136
 Db 133 CGAGACGATCGAGCGCGCTCTACCGGCGACGCTCATGCGCTACGAGCGGCTCTCGCG 192

QY 137 ValLeuIleLeuAsnArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHis 156
 Db 193 GTCTCATCTCTGCTGCTGTCTGTAGCACAGACGCTCAAAACGGTCTCCCATATAGACAC 252

QY 157 LeuValGlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeu 176
 Db 253 GTGGTCGAGCGGTGATTTATGACCCGAGAAAGAGCCCAAGAAAGTTGCAAGACTTAATTCG 312

QY 177 ProHisAsnMetPheTrpValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyr 196
 Db 313 TCTTCAACAAATATGAGTGGTCCCTGCTGATGTTCTGACACCTGCGACGCGCGCGG 372

QY 197 LeuGlyGlyArgGlnLeuAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThr 216
 Db 373 CGAGAGGGCGCCATCCCGGACAGTGGCCCTAAGTTGCTACTAGAGAGCTGAATGTG 432

QY 217 LeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyr 236
 Db 433 TTTCGAGCAAAATGTGGAGTGTCTTCTACTACGACATGATGATACCCCTGCTAC 492

QY 237 ThrGlnValValThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGln 256
 Db 493 ACCGAGTAGTCACTATGCGAGTGCAGTCACTTCTTGGCTTGCCTCATCGGCGCTGAG 552

QY 257 PheLeuAsnProAlaLysAlaTyrProGlyHisGlnLeuAsnProLeuValProValPh 276
 Db 553 TTCTCAGACCTTCAACAGGCGTACAAAGACACCCCTGACCTATGCGTACCCATCTT 612

QY 276 eThrPheLeuGlnPhePhePheTyrValGlyTyrLeuLysValGlyLeuSerArgAlaLe 296
 Db 613 CACCTGTGTCAGATCTTCTTCTACGCTGCGTGAAGTA----- 655

QY 296 ValGlyTyrPargHisGlyGlnArgGlyHisGlyGlnLeuLeuGlnTyrThrArgMetG1 316
 Db 656 -----GCACAGCAGCTTAT----- 670

QY 316 nCysGlnGlnArgLysValSerArgValGlnSerSerGlnAlaTyrTrpArgThrProVal 336
 Db 670 ----- 670

QY 336 IleProAlaThrArgGlnAlaGlnAlaGlyGlnSerLeuGlnProGlyArgArgArgLe 356
 Db 670 ----- 670

QY 356 uTrpTrpIleSerSerSerThrProLeuGlnArgMetMetIleLeuArgProTh 376
 Db 671 -----AACCTTGGAGAGACCAAGACGACCTTGGACCAAT 708

QY 376 rGlyLeuSerThrGlyIleCysArgCysProCysTyrPleuTyrMetArgCysThrArgTh 396
 Db 709 TTC-TTATTTGACCAAAATTCAGAGTGTCTATGCTAGCTGTAGATGAGATGTATGACGAC 767

QY 396 rCysLeuGlyTyrSerArg 402
 Db 768 TGGCGAGGCTGCGCAAAAG 786

RESULT 9
 B0345562 525 bp mRNA linear EST 20-MAY-2002
 LOCUS PM3-NT0312-190501-002-a06 NT0312 Homo sapiens cDNA, mRNA sequence.
 DEFINITION B0345562
 ACCESSION B0345562.1 GI:21009623
 VERSION B0345562.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 525)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovsky-Almeida, S., Britones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

TITLE Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?rl=PM3&rl2=PM3-NT0312-
190501-002-806&rl3=2001-05-19&rl4=1)
Seq primer: puc 18 forward
High quality sequence start: 28
High quality sequence stop: 441.

FEATURES
source
1..525
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NT0312"
/dev_stage="Adult"
/note="Organ: nervous tumor; Vector: puc18; Site:1: Smai;
Site:2: Smai; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 104 a 140 c 164 g 117 t
ORIGIN

Alignment Scores:

Pred. No.: 1.36e-55 Length: 525
Score: 746.00 Matches: 155
Percent Similarity: 91.81% Conservative: 2
Best Local Similarity: 90.64% Mismatches: 9
Query Match: 31.74% Indels: 6
Gaps: 0

US-09-622-964-5 (1-435) x BQ345562 (1-525)

QY 174 LeuSerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsnLeuSerMet 193
DB 23 CTGAGCCTACACACAAACATGTTCTGGGTGGCCCTGGGTGTTTGC-ACCTGTCAATG 81
QY 194 LysAlaTrpLeuGlyGlyArgTlleArg-AspProIleLeuLeuGlnSerLeuAsnG1 213
DB 82 AAGCGGTGCTTGGAGTGGATCTGGGACCCCTATCTCTCCAGACCGTGTGAACGA 141
QY 213 UdeLanPheLeuArg-ThrGlnCysGlyHisLeuTyraLysAspTrpIleSerIleP 233
DB 142 GATGAACACCTTGGCTACCTAGTGTGGACACGTAACCTGACTGGATTAGTATCC 201
QY 233 ToleValTyrrThrGlnValAlaThrValAla-ValTyrrSerPhePheLeuTrnCysLeu 252
DB 202 CAATGGTATACACAGAGTGTGCTGCTGGGTGTACAGCTTCTCCCTGACTTGCTA 261
QY 253 ValGlyArgGlnPheLeuAsnProAlaLysAlaLysAlaTyrrProGlyHisGlnLeuAspLeuVal 272
DB 262 GTTGGGGGGGAGTACAGAAACGACCAAGCCCAACCTGGCCATGAGCTGGACCTGCT 321
QY 273 ValProValPheThrPheLeuGlnPhePhePheTyraGlyTyrrPheLeuTyraGlyLeu 292
DB 322 GTGGCCGCTTCAAGTCTTCTCTTATGTTGGCTGAGGCTGAAGGTGGGCTCC 381
QY 293 SerArgAlaLeuLeuGlyTyrrParGHisGlyGlnArgGlyHisGlyGlnLeuLeuGln 312
DB 382 TCCAGGGCCCTGCTGGGTGGAGGCATGGCCAGAGGGGTCAATGCCAGACGCTGCTGAG 441

QY 313 ThrArgMetGln-CysGln-GluArgLysValSerArgValGlnSerGlnAlaTrpT 332
DB 442 ACGAGGATGACAAAGTGTCTGAGGAAGACGAGCTCTACGGGTGTGAAGACGACGCGTGGC 501
QY 332 rParGThrProValIleProAla 339
DB 502 GCGGCACACCTGTATATCCAGCT 524

RESUT: 10
BM685501
LOCUS
DEFINITION
BM685501
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 526)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
source
1..526
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UI-E-C10-aad-h-10-0-UI"
/clone_lib="UI-E-C10"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pTrf73-Pac (Pharmacia) with a
modified polylinker; Site:1: EcoR I; Site:2: Not I;
UI-E-C10 is a cDNA library containing the following
tissue(s): RPE and Choroid. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pTrf73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dri)18 tail. The
sequence tag for this library is ACCGA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."

Alignment Scores:

Pred. No.: 1.67e-55 Length: 526
Score: 745.00 Matches: 145
Percent Similarity: 67.59% Conservative: 1
Best Local Similarity: 67.13% Mismatches: 0

VERSION BE236901.1 GI:9021619
 KEYWORDS EST.
 SOURCE COW
 ORGANISM Bos taurus

REFERENCE
 AUTHORS
 1 (bases 1 to 508)
 Casas, E., Wray, J.E., White, J., Cho, J., Fahrnkung, S.C., Bennett, G.L., Heaton, M.P., Laegreid, M.W., Rohrer, G.A., Chitko-McKown, C.G., Kelle, J.W., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
 Smith, T.P.L., Grose, N.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
 Sequence evaluation of four pooled-tissue normalized bovine CDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)

TITLE
 JOURNAL
 MEDLINE
 COMMENT

USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross-match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAAACAGCTATGACAT
 BACKWARD: GTTTCACAGTACAGCAGC
 Plate: 56 row: B column: 18
 Seq primer: ATTGAGTACACTATAC.
 Location/Qualifiers
 1..508
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 /clone_1lb="MARC 4BOY"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6, site_1: NotI; site_2: SalI;
 Library made from pooled tissue from day 20 and day 40
 embryos."

BASE COUNT 112 a 144 c 134 g 118 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,15e-51 Length: 508
 Score: 698.00 Matches: 138
 Percent Similarity: 63.36% Conservative: 9
 Best Local Similarity: 59.48% Mismatches: 16
 Query Match: 29.70% Gaps: 2

DB: 10

US-09-622-964-5 (1-435) x BE236901 (1-508)

QY 151 PheProSeRaLaGInHIsLeuValGInaLaGlyPheMetThrProAlaGluHIsLysGln 170
 DB 14 TTTCCAGCCACAGCAGCTGCTGAAGCAGCTTCATGACACCTCGGAACACACAGAC 73
 QY 171 LeuGluLysLeuSerLeuProHIsAsnMetPheTrpValProTrpValTrpPheAlaAsn 190
 DB 74 TTACAAAAAAGCTTACACACAACTCGTTCTGTGATGCCCTGGGTGCTTGGCCAAC 133
 QY 191 LeuSerMetLysAlaTrpLeuGluLysLysLeuGluGluSerLeu 210
 DB 134 TTGTCAACAAAGGATGATGAGCTGATCCGGGACCTATCTGCTCCAGAGCCTG 193
 QY 211 LeuAsnGluMetAsnThrLeuAlaGTrhGInCysGlnHIsLeuTyAlaTyAspTrpIle 230
 DB 194 CTCACGAATGAACACCTTCTGCTACACAGTGTGACAGCTGATGCTGCGACTGATC 253
 QY 231 SerIleProLeuValTyTrhGlnValAlaValAlaValAlaValTySerPhePheLeuThr 250
 DB 254 AGTGTCCCGCTGTGTACACTCAGCTGCTGATGACCTGCGCATATACAGCTTCTCTGCGCT 313

QY 251 CysLeuValGlyArgGlnPheLeuAsnProAlaLysAlaTyProGlyHIsGluLeuAsp 270
 DB 314 TCCCTGATTTGACCGGCACTTTCTGTGAACCCACAGCCCTACCCGGCCACGATATGAC 373
 QY 271 LeuValValProValPheThrPheLeuGlnPhePheTrpValGlyTrpLeuLysVal 290
 DB 374 CTCGTGTGACCCCTTCTGACCTTCTGACGTGCTTCTTCTTATGCGCGC----- 421
 QY 291 GlyLeuSerAlaGlnAlaLeuGluGlyTrpArgHIsGlyLnaArgLysGlnGlnLeu 310
 DB 421 ----- 421
 QY 311 LeuGluThrArgMetGlnCysGlnGluArgLysValSerArgValGluSerGlnAla 330
 DB 421 ----- 421
 QY 331 TrpTrp-ArgTrpProValIleProAlaThrArgGluAlaGluAlaGlyGluSerLeuG 350
 DB 422 TGGTGGAAAG----- 431
 QY 350 UPtGlyArgArgArgLeuTrpTrpGlnSerSerSerTrpProLeuGluArgMetne 370
 DB 432 -----TGCAGATCAGCTCATCAACCATTTGGAGAGATGAC 469
 QY 370 tMetIleLeuArgProThrGlyLeuSerThrGly 381
 DB 470 GATGACTCTCATACCACTGATGTGCACACAGA 503

RESULT 13
 BM562042
 LOCUS
 DEFINITION AGENCOURT.6561919 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5745408
 ACCESSION BM562042
 VERSION BM562042.1 GI:18807794
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NIH-MGC http://mhc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 http://image.llnl.gov
 Plate: IMAGE2768 row: f column: 01
 High quality sequence stop: 634.
 Location/Qualifiers
 1..1067
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="NIH_MGC_118"
 /tissue_type="leucocyte"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6, site_1: NotI; site_2: EcoRV
 (destroyed); RNA source: leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."

FEATURES
 source

BASE COUNT 253 a 303 c 295 g 213 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 8.57e-49 Length: 1067
 Score: 674.00 Matches: 127
 Percent Similarity: 94.81% Conservative: 1
 Best Local Similarity: 94.07% Mismatches: 7
 Query Match: 28.68% Indels: 0
 DB: 13 Gaps: 0

US-09-622-964-5 (1-435) x BM562042 (1-1067)

QY 87 ThrLeuValValThrArgTrpPasnGlnTrgUnsnLeuProTrpProAspArgLeu 106
 Db 11 TCGCTGCTGTCGACCCGCTGCTGACACAGACAGAACTGCGTGGCCGACCCGCTC 70
 QY 107 MetSerLeuValSerGlyPheValGluGlyLysAspGluGlnGlyArgLeuArgArg 126
 Db 71 ATGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 130
 QY 127 ThrLeuValArgTrpAlaAsnLeuGlyAsnValLeuLeuLeuArgSerValSerThrAla 146
 Db 131 AGCTTCATCCGCTGACGCAACCTGCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 190
 QY 147 ValTrpLysArgPheProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAla 166
 Db 191 GTCACAGAGCGCTTCCGACGCGCCGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
 QY 167 GlnHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetPheTrpValProTrpVal 186
 Db 251 GAACACAGACAGTTGGAGAACTGACCTTACACACACATGTTGCTGCTGCTGCTGCTG 310
 QY 187 TrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyGlyArgLysArgAspProLeu 206
 Db 311 TGGTTTCCACACCTGCAATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 370
 QY 207 LeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGlnCys 221
 Db 371 CTCACAGCGCTGCTGACAGCTGACCCACCTGACACAGAGGCTGC 415

RESULT 14
 AL627604 722 bp mRNA linear EST 02-NOV-2001
 LOCUS AL627604 XGC-gastrula silurana tropicalis cdna clone Tgas025m15 5',
 DEFINITION mRNA sequence.
 ACCESSION AL627604.1 GI:16597087
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 western clawed frog.
 Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 722)
 Huckle, E., Taylor, R., Ashurst, J. L., Zorn, A. M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (10_2001)
 Unpublished (2001)
 Contact: Huckle E
 Sanger Centre
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: tropesanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: Tgas025m15.sp6
 Sequencing primer: Sp6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 Location/Qualifiers
 1..722

FEATURES
 source
 /organism="Silurana tropicalis"
 /db_xref="taxon:8364"
 /clone="Tgas025m15"
 /clone_lib="XGC-gastrula"

/dev_stage="gastrula (stages 10.5-13 mixed)"
 /lab_host="Escherichia coli XL1-blue"
 /note="vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dT primed from 5ug of poly A+ RNA from stages
 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
 into pCS107 with EcoRI at the 5' end and NotI at the 3'
 end."

BASE COUNT 156 a 177 c 190 g 196 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 5.25e-49 Length: 722
 Score: 673.50 Matches: 132
 Percent Similarity: 72.32% Conservative: 30
 Best Local Similarity: 58.93% Mismatches: 62
 Query Match: 28.66% Indels: 1
 DB: 9 Gaps: 0

US-09-622-964-5 (1-435) x AL627604 (1-722)

QY 1 MethTrpLeuValValThrArgTrpPasnGlnTrgUnsnLeuProTrpProAspArgLeu 20
 Db 51 ATGACAGTACAGTACAGCAGCAGAGTGCAGAAATCCGCTTGTGAGCTG 110
 QY 21 LeuLeuCystrPargGlySerIleTrpLysLeuLeuTrpGlyLubPheLeuLeu 40
 Db 111 TCGCTGCTATGAGAGGAGGAGCATAATTAATTAATTAATTAATTAATTAATTAATTA 170
 QY 41 LeuCystrTrpTrpIleLeuArgPheIleTrpArgLeuAlaLeuThrGlnGlnGlnLeu 60
 Db 171 CTCATGATCTGCGCTGCTGATCATATACAGTTCTTCTAATGAGACAGAACTT 230
 QY 61 MetPheGluLysLeuThrLeuTrpCysAspSerTrpIleGlnLeuIleProIleSerPhe 80
 Db 231 TACTTTCAGCAAGATGGCCATTACCTATTAATTAATTAATTAATTAATTAATTAATTA 290
 QY 81 ValLeuGlnPheTrpValThrLeuValValThrArgTrpPasnGlnTrgUnsnLeu 100
 Db 291 GTCGTGGTTTATGTCACACCTGCTGTAATTCCTGCTGTAATTCCTGCTGTAATTCCTG 350
 QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlyLysAspGln 120
 Db 351 CCTTCCCGACCGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 410
 QY 121 GlyArgLeuLeuArgArgTrpLeuLeuArgTrpAlaAsnLeuGlyAsnValLeuLeu 140
 Db 411 GGGCGCTGTACCGCGGACCTTATGCGTACCTGACCTATGACGCTATGATTCG 470
 QY 141 ArgSerValSerThrAlaValTrpLysArgPheProSerAlaGlnHisLeuValGlnAla 160
 Db 471 CGCTGTGTTATACCGCTGCTGCTTCAAAAGATTCCTTCAATATGACCATGTTGTAAGAC 530
 QY 161 GlyPheMetThrProAlaGlnHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet 180
 Db 531 GGGTTCATGATCCCTCGAGAGAAAGAACTTGTAGAACTCCAGTCATCTTAAACAG 590
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
 Db 591 TACTGNTTCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgTrpGln 220
 Db 651 ATCCGAGATACCA-TTCCTTTAAATGCTGATGAGAGACTGAAACCTTCCGGGGGAAAC 709
 QY 221 CysGlyHisLeu 224
 Db 710 TGCNGCATGCTC 721

RESULT 15
 BU069061 628 bp mRNA linear EST 11-DEC-2001
 LOCUS BU069061 NIBB Mochii normalized Xenopus tailbud library Xenopus
 DEFINITION laevis cdna clone XL051K03 5', mRNA sequence.

us-09-622-964-5.rst

Page 13

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 628)	Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.	Expressed genes in <i>X. laevis</i> embryo	Unpublished (2001)	Contact: Tadasu Shin-i

TITLE	Expressed genes in X. laevis embryo
JOURNAL	Unpublished (2001)
COMMENT	Contact: Tadashi Shln-I Center For Genetic Resource Information National Institute of Genetics 111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshln@genes.nig.ac.jp
FEATURES	Location/Qualifiers
source	1. 628

BASE COUNT	(McIntire/CNC Institute).
146 a	141 c
161 g	180 t
ORIGIN	

Alignment Scores:	
Pred. No.:	4,36e-48
Score:	662.00
Percent Similarity:	76.92%
Best Local Similarity:	63.08%
Query Match:	28.17%
OB:	13
	13
	628
Length:	123
Matches:	27
Mismatches:	45
Indels:	0
Gapc:	0

022 504-3 (1-435) x BJ069061 (1-628)

Qy MethrIleuHrIyThrSergInValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db ATGAGCGGTACCTACACGCGACGAGTGCAAATGCCGTTTGGGGCTTTATATAACG 31
Qy LeuLeuCySTrPAAGlySerIleTyLysLeuLeuTyGlyGluPheLeuIlePheLeu 40
Db TTTGCGCATATGAGAGGGAGATATATACAACTCTGTACAGAANAATCTGGCTTTTAA 15
Qy LeuCySTyTyTTrIleIleArgPheIleTyArgLeuAlaLeuThrGluGlnGlnLeu 60
Db ATCATGTACCTGGGGGCTCAGCATCATATACAGGTTTTTCTGTAAGAAGAACAAGACTT 21
Qy MethLeuIuLysLeuThrIleTyCyAspSerTyTTrIleGlnIleIleProIleSerPhe 80
Db TTTTCTTGAAGAAAGTGGCCATTTTACTCATTACTATGCAAACTTATCCGTGTCCTTT 212
Qy ValLeuGlyPheTyTTrValThrLeuValValThrArgTTrPAsnGlnTTrGlyuAsnLeu 100
Db GTGTGGGTTTATATATCATCCCTGGGTGAATTCGCTGGTGAACCACTACTCTGTATAT 272
Qy ProTTrProAspArgLeuMetSerLeuValSerGlyPheValGlyGlyLysAspGluGln 120
Db CTTTCCAGACCGCTGTCTATGTGTCTGTCTCTCCGGTACTTTCATGCGGTCTGATGAAG 332
Qy GlyArgLeuLeuArgArgThrLeuIleArgTyTTrAlaAsnLeuGlyAsnValValLeuIleu 140
Db TTTTCTTGTAT 121

Db	392	GGAGGTCTACCGCCGCGCACTGATAGCACTACTGACGAGCTATTCAGAGCACTTTAATCTCG	451
Oy	141	ArgsetValSerThrAlaValTyrTyrAspArgProSerAlaGlnHisLeuValGlnAla	166
Db	452	CGTTCGGTAGTACCCCTGCTCTTCAAAGATTTCTTACATAGACACATGTTGTGAAGCT	511
Oy	161	GLYPheMetThrProAlaGlnHisLysGlnLeuGluTyrLysLeuSerLeuProHisAsnMet	186
Db	512	GGGTCATGACTCGCTCGCTGACAGAGAAAGATTGGAACCTTCAGTGGCTCTCTATTAACAAG	571
Oy	181	PheTrpValProTrpValTrpPheAlaAsnLeuSerMetTyrAla	195
Db	572	TACTGGGTCCGTCGTCGTGGTTCGTGAACCTTGGACGACACAGCA	616

Search completed: July 27, 2003, 19:20:27
Job time : 1744.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model
Run on: July 27, 2003, 14:57:56 ; Search time 258.065 Seconds
(without alignments)
3477.440 Million cell updates/sec

Title: US-09-622-964-5
Perfect score: 2350
Sequence: 1 MTITYSOVANARAGSFSRL.....QLPPSSVEPPLMAPSTSA 435

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1439767 seqs, 1031500376 residues
Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framer.p2n model -DEV=nlh
-O=/cg2_1/USPTO.spool/US09622964/runat.22072003.101120.26142/app-query.fasta_1.2069
-DB=Published_Applications_NA -OFMT=fastap -SUFFIX=tmpp -MINMATCH=0.1
-LOOPT=0 -LOOPTXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-MAXLEN=2000000000 -USER=US09622964 -CCCN=1.131 -runat.22072003.101120.26142
-NCPU=6 -ICPU=3 -NO MAP -LANG=EDUCERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOC -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published_Applications_NA:*

- 1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 5: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 12: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 14: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 15: /cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cg2_6/ptodata/2/pubpna/US60_NEW_PUBCOMB.seq:*
- 17: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	569.5	24.2	1292 15	US-10-198-846-11070 Sequence 11070, A

2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45																	
462	432	333.5	310	300	270	238	194.5	178	177.5	177.5	177.5	176.5	175.5	175.5	175.5	175.5	174.5	174	174	172	172	172	171.5	171	171	171	171	170	170	169.5	169.5	169.5	168.5	168.5	168.5	168.5	168.5	168.5	168.5	168.5	168.5	168.5	167.5																	
19.7	18.4	14.2	13.2	12.8	11.5	10.1	8.3	7.6	7.6	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.4	7.4	7.4	7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.1																	
1350	1263	748	578	462	415	374	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282	282																
US-09-768-826-16	US-09-746-783-3	US-10-027-632-134529	US-10-198-846-11346	US-10-027-632-134530	US-10-198-846-12561	US-09-764-872-615	US-09-880-107-3366	US-09-801-8768-3	US-10-254-669-3	US-10-198-846-1862	US-10-027-632-97877	US-09-969-708-79	US-09-954-456-2116	US-10-027-632-162774	US-10-027-632-153828	US-09-918-995-11849	US-10-027-632-277698	US-09-263-959-1	US-10-081-327-32	US-09-780-172-18	US-09-764-868-1454	US-09-764-868-1455	US-09-764-868-1461	US-10-027-632-256610	US-10-027-632-256611	US-10-027-632-255187	US-10-027-632-112324	US-10-027-632-112325	US-10-027-632-98369	US-10-027-632-110112	US-10-027-632-110113	US-09-764-891-9065	US-09-764-891-7028	US-10-222-334-7	US-09-764-891-6868	US-10-027-632-144268	Sequence 16, App1	Sequence 3, App1	Sequence 134529	Sequence 11346, A	Sequence 124530	Sequence 2361, Ap	Sequence 7843, Ap	Sequence 195, App	Sequence 365, App	Sequence 3, App1	Sequence 1862, Ap	Sequence 97877, A	Sequence 219, Ap	Sequence 162774, Ap	Sequence 153828, Sequence 277698	Sequence 11849, A	Sequence 256610, Sequence 256611	Sequence 255187, Sequence 112324	Sequence 112325, Sequence 98369, A	Sequence 110112, Sequence 110113	Sequence 110114, Sequence 9065, Ap	Sequence 7028, Ap	Sequence 6868, Ap	Sequence 144268,

ALIGNMENTS

RESULT 1
US-10-198-846-11070
Sequence 11070, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lilly, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
TITLE OF INVENTION: STEINMANN, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: THERAPY OF BREAST CANCER
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
 OTHER INFORMATION: n - A,T,C or G
 US-10-198-846-11070

Alignment Scores:

Pred. No.:	1-356-45	Length:	1292
Score:	569.50	Matches:	132
Percent Similarity:	44.86%	Conservative:	25
Best Local Similarity:	37.71%	Mismatches:	69
Query Match:	24.23%	Indels:	124
	15	Gaps:	3

US-09-622-964-5 (1-435) x US-10-198-846-11070 (1-1292)

```

OY 82 LeuGlyPheTYrValThrLeuValValThrArgTTPRPSnGlnTYrGlnLeuPro 101
Db 317 ATAGGTTTATGTTACTCTGTGTAGTGAACCGATGTGGAAACAGCTTGATTTGCC 376
OY 102 TTPRPSAArgLeuMetSerLeuValSerGlyPheValGlnGlyArgGlnGly 121
Db 377 TGGCAGACAGGCTAATGTTCTCATCTCTAGACAGTTCACGAGACGAGACAGG 436
OY 122 ArgLeuLeuArgThrLeuLeuLeuArgTYrAlaAsnLeuGlyAsnValLeuLeuArg 141
Db 437 CGCTGCTTAGAAGACGCTGATGCGCTACGTCATCTCACCTGCTGCTGCTGCTG 496
OY 142 SerValSerThrAlaValTYrArgPheProSerAlaGlnHisLeuValGlnAla 161
Db 497 TCGGTGACACTGCTGTGACAAAGATTCCACAAATGACACAGACAGGCTGGAAG 556
OY 162 PheMetThrProAlaGlnHisGlnLeuGlnLeuGlnLeuSerLeuProHisAsnMetPhe 181
Db 557 TTTATGACACAGATGAAAGAAATTAATTAACACCTCAAGTCTCATCTGAAATAT 616
OY 182 TTPValProTTPRValTTPRPhelaAsnLeuSerMetLysAlaTTPLeuGlyArg 201
Db 617 TGGGTTCATCATCTGTTGGAAATCTTGCAACTAAAGCCGAAATGAAGTGAATC 676
OY 202 ArgASPProIleLeuLeuGlnSerLeuLeuAsnGlnLeuMetAsnThrLeuArgThrGlnCys 221
Db 677 AGAGACAGTGTGATCTGCAATCATGTGATGACTGAATGAAATCATCATCCCTGTGCTG 736
OY 222 GlnHisLeuTYrAlaTYrAspTTPRILESerIleProLeuValTYrThrGlnValThr 241
Db 737 AGCCTCTATTCGCTGTATGACTGGGTTCCGCTGCTGTTACACCCAGGTA----- 790
OY 242 ValAlaValTYrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnProAla 261
Db 790 ----- 790
OY 262 LysAlaTYrProGlnHisGlnLeuAspLeuValValProValPheThrPheLeuGlnPhe 281
Db 790 ----- 790
OY 282 PhePheTYrValGlyTYrPheLeuLysValGlyLeuSerArgAlaLeuLeuGlyTYrPArgHis 301
Db 790 ----- 790
OY 302 GlyGlnArgGlyHisGlyGlnGlnLeuLeuGlnUTHrArgMetGlnCysGlnGlnArgLys 321
Db 790 ----- 790
OY 322 ValSerArgValGlnSerSerGlnAlaTTPRTPRArgThrProValIleProAlaThrArg 341
Db 790 ----- 790
OY 342 GluAlaGlnAlaGlyGlnSerLeuGlnProGlyArgArgArgLeuTTPRTPRgln-SerSe 361
Db 791 ----- 791
OY 361 rSerSerThrProLeuGlnArgMetMetMetIleLeuArgProThrGlyLeuSerThrGln 381
Db 800 CTTATCAACCTTTGGAGAGATGATGATATTTTGAACCTAACTAGTGTGATGACAGA 859
  
```

```

OY 381 ylleCysArgCysProCysTTPRLeuTrpMetArgCysThrArgThrCysLeuGlyTYrPse 401
Db 860 AATTGGACAGTCTCTCTTTAGCTGTGGACGAAATGACACAGACTTACCAGATGAAG 919
OY 401 rArgThrCysThrGly-----IleSerProSerHisSerProProThrGlnLeuLeuPr 419
Db 920 AAGGACATTACTGTGGACGATTCCTGCTGCCACCATACATGACAGCTCTG-- 977
OY 419 oProSerSerValGlnProProLeuLeuTP 428
Db 978 ----ACTATCTGCAATCCCTCATTTCTGG 1001

RESULT 2
US-09-768-826-16
Sequence 16, Application US/09768826
Patent No. US20020012966A1
GENERAL INFORMATION:
APPLICANT: Shi et al.
TITLE OF INVENTION: 18 human secreted proteins
CURRENT APPLICATION NUMBER: US/09/768,826
CURRENT FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: PCT/US00/22350
PRIOR FILING DATE: 2000-08-15
PRIOR APPLICATION NUMBER: 60/148,759
NUMBER OF SEQ. ID NOS: 61
SOFTWARE: PatentIn Ver. 2.0
SEQ. ID NO 16
LENGTH: 1350
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1135)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1209)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1229)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1266)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1285)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1287)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1290)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1295)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1305)
OTHER INFORMATION: n equals a,t,g, or c
  
```



```

NAME/KEY: SITE
LOCATION: (1324)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16

```

```

Alignment Scores:
Pred. No.: 3 82e-35
Score: 462.00
Percent Similarity: 46.268
Best Local Similarity: 41.288
Query Match: 19.664
DB: 10
Gaps: 6

```

```

US-09-622-964-5 (1-435) x US-09-768-826-16 (1-1350)
QY 172 GULYSLEUSERLEUPROHISAMETHPYVALPROTPYVALTPRPHEALASNULEU 191
DB 38 GAGACCTGAAATCCGACTTCAACAGTACGTGGCCCTCGCTGTTTACCAACCTG 97
QY 192 SERMETLSALATPRLEUGLYLYARGYLEARGSPROILLEULEUGINSEULEU 211
DB 98 GCGGCCGAGCCGAGGAGGAGGCGGGAATACGTACATATCGCTCTCTCTCTCTCTG 157
QY 212 ASNGUMETLANTHLEUARGTHGNCYSGLYHISLEURYALATYRASPPTILSEU 231
DB 158 GAGAGGCTGAAACAGTACCGGCAAGTGCATGCTATTCCACTGTGATGATCAGC 217
QY 232 ILEPROLEUVALTYRTHGVALTHVALTHVALTHVALTHVALTHVALTHVALTH 251
DB 218 ATCCCTCGCTACACCAAGTGGTACCATAGCCGCTCTCTCTCTCTCTCTCTCTG 277
QY 252 LEUVALGLYARGINPHELEUASPRO-----ALALYSALATYR----- 264
DB 278 CTGTTGGCCGCGAGTTTGGAGCCAGGCGGCTGCCAACCCTCAGAACCTCTG 337
QY 265 ---PROGLYHISGLU-----LEUASPLEUVALPROVALPHE 276
DB 338 AACCCAGCCAGAGCCAGCCCGGAGACCCGAGACATGACGTGCTCTCTCTCTC 397
QY 277 THPHEUGLINPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 296
DB 398 ACTCTCTCTGAGTTCTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449
QY 297 LEUGLYTPARGHISGLYINARGGLYHISGLYINLEUGLINLEUGLINUTHR 316
DB 449 ----- 449
QY 317 CYSGLINGLYARGLYVALSERARGVALGLUSERSERGINALATPTPTPTPTPT 336
DB 449 ----- 449
QY 337 ILEPROALATHARGGLUALAGLUALAGLYUSERLEUGLUPROGLYARGARGLEU 356
DB 449 ----- 449
QY 357 TRTPRGLINSEUSERSERTHPROLEUGLUNARGMETWETHTTILEUARGPROTHR 376
DB 450 -----TATCAACCCATTGGTGAGATGATGACGACTTGGAGCAAAATC 494

```

```

QY 377 GLYLEUSERTHRGILYILECYARGCYSPTCYSTPLPMTETARGCYTHARGTHR 396
DB 495 AGCTCATAGACCCGACACTTGGAGGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 554
QY 397 CYSLEUGLYTPRSPERARGTHRCYSTRGILYILESERPROSERHISERPROPTH 416
DB 555 TTCCCCCGCTGAGAGGACCACTGCTGAGTGGAGGACCCGAGCCACCTTACA 611
QY 417 LEULEUPRO-----PROSERVALGLUPROPROLEUTPRALAPROPROSERTH 434
DB 612 CTGTCGCGCAGCGGCGGAGTCTCTGCGGCTCATTCCTGAGGCTCACCCTCA 671
QY 435 ALA 435
DB 672 GCA 674

```

RESULT 3

```

US-09-746-783-3
Sequence 3, Application US/09746783
Publication No. US20030044935A1
GENERAL INFORMATION:

```

```

APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavallee, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.
Howes, Steven H.
Fechtel, Kim

```

```

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MLASINGLO, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214

```

```

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

```

```

Alignment Scores:
Pred. No.: 2 82e-32
Score: 432.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 18.38%
DB: 12
Gaps: 0

```

US-09-622-964-5 (1-435) x US-09-746-783-3 (1-1263)

OY 360 SerSerSerThrProLeuGluArgMetMetIleLeuAlaProThrGlyLeuSer 379
Db 2 AGCGGCTATCAACCCCTTTGGAGAGATGATGATTTTGGAGCACTGATTTGG 61
OY 380 ThrGlyIleCysArgCysProGlyTrpLeuTrpMetLeuArgCysThrArgGlyLeuGly 399
Db 62 ACAGGAATTTGCAGAGTCTCCCTGCTGTGTGATGATGACACAGAGAGCTGCTCGA 121
OY 400 TrpSerArgThrCysThrGlyIleSerProSerHisSerProProThrGlyLeuLeuPro 419
Db 122 TGGAGCCGGACATGATCTGGAATTAAGCCGAGCCAGCAAGCCCTTACAGCTGCTCCG 181
OY 420 ProSerSerValGluProProLeuTrpAlaProProSerThrSerAla 435
Db 182 CCCAGTTCGGTCGAGCCCTCTTTATGGGCTCCACCTTCACATCAGCC 229

RESULT 4

US-10-027-632-134529/c
; Sequence 134529, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 134529
; LENGTH: 748
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134529

Alignment Scores:

Pred. No.: 4, 77e-23 Length: 748
Score: 333.50 Matches: 83
Percent Similarity: 38.188 Conservative: 1
Best Local Similarity: 37.738 Mismatches: 6
Query Match: 14,198 Indels: 131
DB: 15 Gaps: 1

US-09-622-964-5 (1-435) x US-10-027-632-134529 (1-748)

OY 152 ProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLeu 171
Db 712 CCCACCGCTTTTTCACCTCCACTCTGACGCTTTATGACTCGGAGAACACACAGCTG- 654
OY 172 GluLysLeuSerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsnLeu 191
Db 653 GAGAAACTGACCTACACACACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTG 594
OY 192 SerMetLysAlaIlePheGlyGlyArgIleArgAspProIleLeuGlnSerLeu 211
Db 593 TCATGAAAGCGCTGCTGGAGTGAATCCGGGACCTTAACCTGCTCCAGAGCTGCTG 534

OY 212 Asn----- 212

Db 533 AACGTGAGCCACTGTACAGACAGAGGCTGCCCGCAGAGTGGAAAGGTTGTGTCCACAGG 474

OY 212 ----- 212

Db 473 AAACAGGTTCTTCAAAAGAGAGCCTTGCGCCCTGAGAGTCTTCCGAGACCGGAGG 414

OY 212 ----- 212

Db 413 TGGGTTGAGAAATCTTTCCACAGCAATCCACAGCCGAGGTGCTCTTCAGAGG 354

OY 212 ----- 212

Db 353 CCCCTCCCTCTTCAGAGTGTGTGAGTCTGTGTTCCCTTTGATAGTAGAAGCTG 294

OY 212 ----- 212

Db 293 AGACAAAGAGGTTAGTACCTTCCATGGCCACACAGCCAGAAATGACCATAGGTA 234

OY 212 ----- 212

Db 233 CCAGCCCTGTACTGAKAAGAGTGGGGCGAGCCAGAGGTGGGGGAGGTGTGT 174

OY 213 ----- 212

Db 173 CAGAACCCCATCCCTCTTCTGCCCCCAGAGATGAACACCTTGCTACTGATGTG 114

OY 222 yHsLeuTrpAlaTrpAspTrpIleSerIleProLeuValTrpGlnValTrp 241

Db 113 ACACCTGTATGCTTACGACTGATGATATTCACCTGTGTATACAGAGTACAGACT 56

RESULT 5

US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198, 846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306, 220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11346
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
; LOCATION: 43, 50, 51, 54, 57, 853
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346

Alignment Scores:

Pred. No.: 1, 09e-20 Length: 853
Score: 310.00 Matches: 57
Percent Similarity: 83.338 Conservative: 8
Best Local Similarity: 73.088 Mismatches: 13
Query Match: 13,198 Indels: 0
DB: 15 Gaps: 0

US-09-622-964-5 (1-435) x US-10-198-846-11346 (1-853)

OY 83 GlyPheTrpValTrpLeuValValThrArgTrpAsnGlnTrpGlnAsnLeuProTrp 102

```
Db 273 GGGTTTATGTTACTCTGTAGTGAACCGATGCGAACCAGTTTGGAAATTTCCCTCG 332
QY 103 ProAsparGLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGlnGlyArg 122
Db 333 CCAGACAGGCTAATGTTCTCATCTCTAGCGATGTCAGCAACCGACGACGACGCGCGC 392
QY 123 LeuLeuArgArgThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 142
Db 393 CTGCTTAAAGGAGCGGATGCGCTACGTCATCTCAGCTCCCTGCTCATCTTCCGCTCG 452
QY 143 ValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
Db 453 GTGAGCACTGCTGTGTACAAAAGATTTCACCAATGAGCAACGCTGTGTAAGCA 506

RESULT 6
US-10-027-632-134530/c
; Sequence 134530, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218, 006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198, 676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193, 483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185, 218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167, 363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156, 358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146, 002
; NUMBER OF SEQ ID NOS: 1999-08-09
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134530

Alignment Scores:
Pred. No.: 5.86e-20
Score: 300.00 Length: 578
Percent Similarity: 93.44% Matches: 56
Best Local Similarity: 91.80% Conservative: 1
Query Match: 12.77% Mismatches: 2
DB: 15 Gaps: 1

US-09-622-964-5 (1-435) x US-10-027-632-134530 (1-578)

QY 152 ProSerAlaGlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLysGlnLeu 171
Db 201 CCTCTTCACTCCAC-----TCTGAGCGCTTTATGACCTCCGAGAACACACAGAGTTG 148
QY 172 GluLysLeuSerLeuProHisAsnMetPheTyrValProTyrValTyrPheAlaAsnLeu 191
Db 147 GAGAAACTGAGCTTACACACACAAACATGTTCTGTGGTCCCTGGGTGTGTTGCCAACCTG 88
QY 192 SerMetLysAlaTyrLeuGlyLysArgPheProSerAlaGlnHisLeuValGlnSerLeu 211
Db 87 TCAATGAAAGCGCTGCTGTGAGGTGTAATCCGGGAGCCTATCTGCTCAGAGCTCTG 28
QY 212 Asn 212
Db 27 AAC 25

RESULT 7
US-10-198-846-2561/c
; Sequence 2561, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; FILE REFERENCE: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198, 846
; PRIORITY FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306, 220
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2561
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 390..462
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561

Alignment Scores:
Pred. No.: 3.46e-17
Score: 270.00 Length: 462
Percent Similarity: 83.58% Matches: 50
Best Local Similarity: 74.63% Conservative: 6
Query Match: 11.49% Mismatches: 11
DB: 15 Gaps: 0

US-09-622-964-5 (1-435) x US-10-198-846-2561 (1-462)

QY 82 LeuGlyPheTyrValThrLeuValThrArgTyrPheAsnGlnTyrGluAsnLeuPro 101
Db 256 ATAGGTTTATGTTACTCTGTAGTGAACCGATGCGAACCAGTTTGGCAATTTGCC 197
QY 102 TyrProAsparGLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGlnGly 121
Db 196 TGGCCAGACAGCTAATGTTCTCATCTCTAGCGATGTTTACGAGACGACGACGCGG 137
QY 122 ArgLeuLeuArgArgThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 141
Db 136 GCGCTGCTTAAAGGAGCGCTGATGCGCTACGTCATCTCAGCTCCCTGCTCATCTTCCG 77
QY 142 SerValSerThrAlaValTyr 148
Db 76 TCGGTGAGCACTGCTGTGTAC 56

RESULT 8
US-10-198-846-7843
; Sequence 7843, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; FILE REFERENCE: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198, 846
; PRIORITY FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306, 220
; NUMBER OF SEQ ID NOS: 14084
```

SOFTWARE: FastSeq for Windows Version 4.0
Seq ID NO 7843
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-10-198-846-7843

Alignment Scores:
Pred. No.: 5, 2e-17
Score: 270.00
Percent Similarity: 83.58%
Best Local Similarity: 74.63%
Query Match: 11.49%
Length: 615
Matches: 50
Conservative: 6
Mismatch: 11
Indels: 0
Gaps: 0

US-09-622-964-5 (1-435) x US-10-198-846-7843 (1-615)

OY 82 LeuGlyPheTyrValThLeuValThrArgTyrPheAsnGlnTyrGluAsnLeuPro 101
DB 352 ATAGCGTTTATGACTCTGCTAGTGAACCGATGTGGAACCGATTGGAATTGCCC 411
OY 102 TTPProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGlnGly 121
DB 412 TGGCCAGACAGCGTAAGTTCCTATCTACAGCTGTTCACGGAAGCGACGACGCG 471
OY 122 ArgLeuArgArgTyrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeuArg 141
DB 472 CGCTGCTTGAAGAGCGTGTGACGCTACGTCATCTACATCCCTGCTCATCTTCCG 531
OY 142 SerValSerThrAlaValTyr 148
DB 532 TCGGTGACGACACTGCTGTGTAC 552

RESULT 9

US-09-764-872-195
Sequence 195, Application US/09764872
Publication NO. US20030050231A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA125
CURRENT APPLICATION NUMBER: US/09/764,872
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 957
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 195
LENGTH: 507
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (490)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-872-195

Alignment Scores:

Pred. No.: 5, 01e-14
Score: 238.00
Percent Similarity: 72.29%
Best Local Similarity: 55.42%
Query Match: 10.13%
Length: 507
Matches: 46
Conservative: 14
Mismatch: 23
Indels: 0
Gaps: 0

US-09-622-964-5 (1-435) x US-09-764-872-195 (1-507)

OY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 37 ATACCGGTTTCATACACTCTCAAGAGGCGGAGCCCGCTTGGAGGTTCTGGCCG 96
OY 21 LeuLeuCysTyrArgGlySerIleTyrLysLeuLeuTyrGlyGluPheLeuIlePheLeu 40
DB 97 CTTTCGCCGCTGAGGAGGAGCATCTACAGCTCTCTACAGGAATTCCTCTCTTGG 156

OY 41 LeuCysTyrTyrIleThrPheIleTyrArgLeuAlaLeuThrGlnGlnGlnLeu 60
DB 157 GCCTGTAGCGCTGTGCTTACATACACCTACGCTGCTGACCCAGGACAGAGTAC 216
OY 61 MetPheGluLysLeuThrTyrCysAspSerTyrIleGlnLeuIlePheSerPhe 80
DB 217 GTGTATGCTCAGGTGCGCGGTACGCAACCGCTCAGCAGACACTTATCCCTGCTT 276
OY 81 ValLeuGly 83
DB 277 GTATTGGCT 285

RESULT 10

US-09-764-872-615
Sequence 615, Application US/09764872
Publication NO. US20030050231A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA125
CURRENT APPLICATION NUMBER: US/09/764,872
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 957
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 615
LENGTH: 13743
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-872-615

Alignment Scores:

Pred. No.: 9, 04e-08
Score: 194.50
Percent Similarity: 41.78%
Best Local Similarity: 28.29%
Query Match: 8.28%
Length: 13743
Matches: 86
Conservative: 41
Mismatch: 86
Indels: 93
Gaps: 17

US-09-622-964-5 (1-435) x US-09-764-872-615 (1-13743)

OY 157 LeuValGlnAlaGlyPheMetThrProAlaGlnLysGlnLeuGlyLysLeuSerLeu 176
DB 5025 CTGTCTCAGTGTGCTGCTAGCAGCAGCATACCTAGTGGGACATCCCTAGC 5081
OY 177 ProHisAsnMetPheTyrValProTyrVal 193
DB 5082 TCCAGATTCTGGGTGGAGTATTCAGGAACAGGTGGAAGGGGACCTACTGGGA 5135
OY 194 LysAlaTyrPheGlyGlyArgIleArgAspPro 207
DB 5136 CGTGCAAGACTGCGCAGGCGCAGAGGACGAGAGGTGGGGGTAGCAGTGTGCTA 5195
OY 208 GlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGlnCysGlyHisLeuTyr 226
DB 5196 CCTTCA 226
OY 227 TyrAspTyr 242
DB 5244 CACCC-TGGACCAAGCCACATCAATTCCT 5293
OY 243 AlaValTyrSerPhe 258
DB 5294 TCCTCACTGGCCGCTTCTGCGCCCTGTCCTTATGCTATTTGTACCCACACACCT 5353
OY 259 AspProAlaLysAlaTyrProGlyHisGlnLeuAspLeuValAlaProValPheThrPhe 278
DB 5354 GATGATGCCAAGATTGA-CCACATCATCTCTTATAAACAATTCACATA 5403
OY 279 LeuGlnPhePhePheTyrVal 285
DB 5404 GTTTCATCTGCTCAAGATTAATCAAGGCCAGGACAGTGGCTACAGCCTATATATCC 5463

Alignment Scores:	
Pred. No.:	0 000119
Score:	177.50
Percent Similarity:	42.47%
Best Local Similarity:	28.77%
Query Match:	7.55%
DB:	11
	1
	Gaped:
	60
	Indels:
	109
	Mismatches:
	40
	Matches:
	84
	Length:
	148567

QY	115	lgluglylvsaspgingllyargleuleuaghythrleuilearytyralasneu	134
		::: :::	
Db	80346	GAAGAAAAGACAAAGTTCACCTAGTATGACGATAGTATACAGCTCCTTAGGTAC	8028
QY	135	GLysnValleuleuIeuarSerValSerthrAlaValItyrlyAsnGlypheProSerAla	154
		::: ::: ::: :::	
Db	80286	ATAAAAAGATATTTAAAAAGGATATTAACACCAACCAAAAGGAAAAAACGACA	80222
QY	155	GLnHisleuValGlnAlaGlyPheMetThrProIaGlu-----HisIysGlnLeuGlu	172
		::: ::: ::: :::	
Db	80226	CAATGCTACTATTTCATTAATTTTTTTTTGAAAAATTAATAGTACCTAAATGCTGGGAT	80167

193 MelLysALATrPLeuGlyGlyVrArgLeuArgAspPro-----IleLeuLeu 207
80128 TTTAAAAGCCCAATAGATGTAAAGAAAATATTCTCTCTACCTCGATATTTCTAAATACAT 80069

[illegible]

270	HebSulHICysLeuValGlyArgGlnPhe-----	257
b		
79966	TTTTGTGATGTTTGTGTAACCTTTCTTTTAAATAAGAAATTTTATGCAATGTTTGC	79907
258	-----LeuAspProAlaIleValTyrProGlnIleLeuAspLeuVal---Val	273
y		
79906	ATGCTACAGAGTTGGTTCAACCAAGGCCGCCA-----CTGCACATTTGTTGATGTTG	79856
b		
274	ProValPheThrPheLeuGln-----PheAspProMet-----	264
y		

[illegible][illegible]

```

QY      355 ArgLeuTrpTrpGlnSerSerSerThrProLeu 366
          ||||| | |
          |||
          || :: |||||
Db      79558 AGGTTG-----CAGTGAGCCGAGATCACACACACTG 79539

```

RESULT 13
US-10-254-869-3/C
; Sequence 3, Application US/10/254869
; Publication No. US20030027307A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF

```

1 CURRENT APPLICATION NUMBER: US/10/254,861
2
3 CURRENT FILING DATE: 2002-09-26
4
5 NUMBER OF SEQ ID NOS: 8
6
7 SOFTWARE: FastSeq for Windows Version 4.0.C
8
9 SEQ ID NO 3
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
102
```

```

? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(148567)
? OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

```

Alignment Scores:
Pred. No.:

Pred. No.:	0.000119	length:	1
Score:	177.50	Matches:	8
Percent Similarity:	42.47%	Conservative:	1
Percent Local Similarity:	28.77%	Mismatches:	1
Query Match:	7.55%	Indels:	6
Query:	15	Gaps:	1

US-09-622-964-5 (1-435) x US-10-254-869-3 (1-148567)

115 GUGUGLYSPAGSGLGGLGlyArgLeuLeuArgThrLeuLeaArgTYRAlaAsnLeu 134
||||| : : : : :
80346 GAAGGAAAAAACAACAACTTACACACTATGTATGACGTAAGTATACAGCGCTTAGGTAC 8028
||||| : : : : :
135 GlysAlaValLeuLeuLeuArgSerValSerThrAlaValTyrAsnArgPheProSerAla 154
: : : : : : : : : : :
80286 ATTAAAGAGATTTATTAAAAAGCTATTAACACACAAACAAAGAAAAAGCACA 80222
: : : : : : : : : : :

[illegible][illegible]

		GAGCAGACCCCAATCTTTTATT	79967
248	PheLeuThrGlySerValGlyArgGlnPhe-----		257
79966	TTTTGCATTGTTTTAAACTTCCTTTTAATAAAGAATTTTATCAATGTTGCC		79907
258	-----LeuAsnProAlaIysAlaIaTyrrProGlyHisGluLeuAspLeuVal---Val		273
	:: ::: :: :: ::		

```

Db 79906 ATGCTAGAGTTGATTCACAAACAGCCGCCCA-----CTGCATGATGATGTG 79856
QY 274 ProvalPheThrPheLeuGln-----PhePhePheThr---- 284
Db 79855 GCAGCTTTCTTATCTACAGCTGCTGACCACTATATCTTTATTTCTAGTAATT 79796
QY 285 -----ValGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeu 297
Db 79795 TTTTGTAAAAAACCAGAGGTGGCGGTGGTGCAG--TGCTCAGCCCTGCAATCCA 79739
QY 298 Gly--TrpArgHisGlyGlnArgGlyHisGlyGlnLeuLeuGlnThrArgMetGln 316
Db 79738 GCACCTTTGGAGGCTGAAGCGGTGGATCAGAGCTGAGAGTTTGAGACCGAGCTTACC 79679
QY 317 CysGlnGluArgLysValSer-----ArgValGluSerSerGlnAlaTrpArgTrp 334
Db 79678 AACATGTTGAACCCGCTCTACTACTAAATAATGACTGGCATGTGGCGCGCA 79619
QY 335 ProValIleProAlaTrpArgGluAlaGlyGlnLeuSerLeuGlnProGlyArgArg 354
Db 79618 CCTATATCTCAGCTACTCAGAGAGCTGAGGTAGAGTAACACTTGAACCTGGGCGCGG 79559
QY 355 ArgLeuTrpTrpGlnSerSerSerSerTrpProLeu 366
Db 79558 AGGTG-----CAGTGCAGCCGAGATCACACACTG 79529

RESULT 14
US-10-198-846-1862
; Sequence 1862, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1862
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 15, 553, 561, 570, 671, 734, 739, 745
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1862

Alignment Scores:
Pred. No.: 7 21e-08 Length: 751
Score: 177.00 Matches: 50
Percent Similarity: 55.96% Conservative: 11
Best Local Similarity: 45.87% Mismatches: 40
Query Match: 7.53% Indels: 10
Gaps: 3

US-09-622-964-5 (1-435) x US-10-198-846-1862 (1-751)
QY 82 LeuGlyPheTrpValThrLeuValValThrArgTrpAspGlnIleuLeuPro 101
Db 372 ATAGGTTTATGTTACTGTGTAGTGAACGATGGTGAACCACTTTGCAATTTGCC 431
QY 102 TrpProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGlnGly 121
Db 432 TG-CCAGACAGGCTAATGTTCTTCATCTTACAGAGTGTTCACGAGGACGACGCG 490

```

```

QY 122 ArgLeuLeuArgTrpThrLeuLeuLeuArgTrpAlaAsnLeuGlnValLeuLeuArg 141
Db 491 CCGCTGCTTAAGAGACGCTGATGCGCTC-GTCAATCTACCTCTTCTTATCTTTCT 549
QY 142 SerValSerThrAlaValTrpLysArgPhe-----ProSerAlaGlnIleuLeu 157
Db 550 CGGAGAGCTGAGTGTCTCTGCGCTTAACTAGTAGGATCCCGGCTGACGAGAAATTC 609
QY 158 ValGlnAlaGlyPheMetTrpProAla-----GlnHisLysGlnLeuLysLeuSer 175
Db 610 ATTAATCTATCAACCTCACTACAGGGGCGGCTCCACTTTTCTTCTTGTAGGGTAATG 669
QY 176 LeuProHisAsn-----MetPheTrp 182
Db 670 CMCCTGCGAAGCTGATAGTGTCTGG 696

RESULT 15
US-10-027-632-97877
; Sequence 97877, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 97877
; LENGTH: 1712
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97877

Alignment Scores:
Pred. No.: 2.6e-07 Length: 1712
Score: 176.50 Matches: 83
Percent Similarity: 41.30% Conservative: 50
Best Local Similarity: 25.78% Mismatches: 73
Query Match: 7.51% Indels: 18
Gaps: 17

US-09-622-964-5 (1-435) x US-10-027-632-97877 (1-1712)
QY 133 AsnLeuGlyAsnValLeuLeuLeuArgSerValSerThrAlaValTrpLysArgPhePro 152
Db 772 TCTTTAAGCTCTCAATATCTCTCTCCACAGGTGTGACATGCTCATGTTTGAAGATATGTT 831
QY 153 SerAlaGln--HisLeuValGlnAlaGlyPheMetTrpProAlaGlnHisLysGlnLeu 171
Db 832 GAAACAGAGACCCATGAGAGAAACACACAGCTTATGTTCCATAGTCACAC----- 882
QY 172 GlnLysLeuSerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsnLeu 191
Db 883 TCACCTCAGCTCTACTACTAT-----GTTATTTTCTTTTAACTTG 921
QY 192 SerMetLysAlaTrpLeuGlyGlyArgLysAspProIleLeuLeuGlnSerLeuLeu 211
Db 922 -----CTTCTTTAGTCTCTATTTCTT 942

```

```
QY 212 AsnGluMetAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIleSer 231
Db 943 -----TATTT-TTGCTTATATTTATATATCT 968
QY 232 IleProLeuValTyrThrGlnValThrValAlaValTyrSerPhePheLeuThrCys 251
Db 969 CCTCCTTTGTTATACAGCATTTTC-----TTTATTCATTA 1007
QY 252 LeuValGlyArgGlnPheLeuAsnProAlaLysAlaTyrProGlyHisGlnLeuAspLeu 271
Db 1008 TTCATTAGCAACAATACATTA-----TATAAATGACTTAA 1046
QY 272 ValValProValPheThrPheLeuGlnPhePhePheTyrValGly----- 286
Db 1047 ATGTATTATTTCTTCGGTGTATATCTCCCAAGTGTACATATCTGTTTCAGTATGCT 1106
QY 287 -----TrpLeuLysValGlyLeuSer 293
Db 1107 AATATGCACATTTTAAATGCTGTGAGGCGACGCGCATGCTC-----ACGCTGTAAATC 1163
QY 294 ArgAlaLeuLeuGlyTyrPargHisGlyGlnArgGlyHisGlyGlnGlnLeu----- 310
Db 1164 CCAGCAGCT-TTGGA-----GGCCGAGCGCAGGAGATCAAGTAAGT 1204
QY 311 -----LeuGluThrArgMetGlnCysGlnGlnArg-----LysValSer 323
Db 1205 TGGGAGTTCGAGACCGCTGACCAACATGGAATAAACCCCGCTCTACTATAAATACAA 1264
QY 324 ArgValGluSerSerGlnAlaThrPargThrProValIleProAlaThrArgGluAla 343
Db 1265 AAAAT-----AGCCAGCATGGTGGCGCATGCTATTAATCCAGCTACTCGGAGGCT 1318
QY 344 GluAlaGlyGluSerLeuGluProGlyArgArgArgLeuTrp----- 357
Db 1319 GAGCCAGGAAATCGCTTGAACTGGGAGCGGAGGTTGCCAGTGAAGCAAGTCGGGCA 1378
QY 358 -----TrpGlnSerSerSerThrProLeuGlnArgMet----- 369
Db 1379 TTGCACCTCCAGCCGACGACAAAGCAAGCAAACTCCGTCTCAAAAAAATAATGCC 1438
QY 370 -----MetMetIleLeuArgProThr-GlyLeuSerThrGlyIleCysArgCysProCys 387
Db 1439 TGTGAGATGAAGATGATCAGAGACCTGACATC-----TG 1474
QY 387 strPleuTrpMetArgCysThrArg-----ThrCysLeuGlyTyrSerArg 402
Db 1475 TGTGTGTATTTTAAATTTTGTAGAGATGGGTCTCCCTGTGTGCCAGGCTGCTCTCA 1534
QY 402 gThr 403
Db 1535 AACT 1538
```

Search completed: July 27, 2003, 20:49:52
Job time : 340.065 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:08:55 ; Search time 53.7174 Seconds

(without alignments)
2483.449 Million cell updates/sec

Title: US-09-622-964-5

Perfect score: 2350

Sequence: 1 MTTTYSQVANARLGSFSL.....QLPPSSVEPLMAPPTSTA 435

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n_model -DEV=x1h
-DB=us09_1/USPTO.spool/US09622964/rnat.22072003.101119.26036/app.query.fasta_1.2069
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09622964.ecn.1.1.79_etunal.22072003.101119.26036 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -FGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Issued.Patents_NA:*
- 2: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/Backfill1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	7.6	2822	4	US-08-819-872A-10
2	178	7.6	2822	4	US-08-819-872A-11
3	178	7.6	2822	4	US-08-819-872A-11
4	166	7.1	10380	4	US-09-077-354B-3
5	166	7.1	246240	2	US-08-724-394A-20
6	166	7.1	246240	2	US-08-724-394A-21
7	166	7.1	246240	2	US-08-724-394A-22
8	165	7.0	5590	4	US-09-050-159-129
9	162.5	6.9	50000	4	US-09-146-053-3
10	161.5	6.9	70000	4	US-09-851-896-3
11	160.5	6.8	3507	1	US-08-832-883-67
12	160.5	6.8	3507	2	US-08-832-877-67

13	158	6.7	176373	3	US-09-128-155-17	Sequence 17, Appl
14	157.5	6.7	3748	2	US-08-958-240-1	Sequence 1, Appl1
15	157	6.7	5037	4	US-09-705-299-13	Sequence 13, Appl
16	157	6.7	8396	4	US-09-328-114A-1	Sequence 1, Appl1
17	157	6.7	8409	4	US-09-167-681-37	Sequence 37, Appl
18	155.5	6.6	5581	4	US-08-973-544-1	Sequence 1, Appl1
19	155.5	6.6	29629	4	US-09-729-995-3	Sequence 3, Appl1
20	154.5	6.6	168575	4	US-09-426-290-1	Sequence 1, Appl1
21	153.5	6.5	112132	4	US-09-741-150-3	Sequence 1, Appl1
22	153	6.5	1524	4	US-09-056-105-2	Sequence 2, Appl1
23	153	6.5	1559	2	US-08-417-174-1	Sequence 1, Appl1
24	153	6.5	1559	2	US-08-231-565A-1	Sequence 1, Appl1
25	153	6.5	1559	2	US-09-007-961-1	Sequence 1, Appl1
26	153	6.5	1559	4	US-09-741-154-3	Sequence 1, Appl1
27	153	6.5	16389	4	US-09-267-439-1	Sequence 1, Appl1
28	153	6.5	99500	4	US-09-798-096-10	Sequence 10, Appl
29	152.5	6.5	162450	4	US-09-145-882-1	Sequence 1, Appl1
30	152	6.5	7152	4	US-09-167-681-29	Sequence 29, Appl
31	151.5	6.4	35080	3	US-08-814-095-7	Sequence 7, Appl1
32	151	6.4	7620	1	US-07-767-135-1	Sequence 1, Appl1
33	151	6.4	7620	1	US-07-841-652-1	Sequence 1, Appl1
34	151	6.4	14636	4	US-09-173-914-6	Sequence 17, Appl
35	151	6.4	15602	4	US-09-844-634-17	Sequence 17, Appl
36	151	6.4	87350	3	US-08-781-891-79	Sequence 3, Appl1
37	151	6.4	87543	4	US-09-791-211-3	Sequence 3, Appl1
38	150.5	6.4	6235	4	US-09-305-384-5	Sequence 5, Appl1
39	150.5	6.4	6679	4	US-09-305-384-1	Sequence 1, Appl1
40	150.5	6.4	84495	4	US-09-797-906-3	Sequence 3, Appl1
41	150	6.4	4136	4	US-09-103-875-2	Sequence 2, Appl1
42	150	6.4	4421	2	US-08-257-963B-9	Sequence 9, Appl1
43	150	6.4	4421	4	US-08-367-841A-6	Sequence 6, Appl1
44	150	6.4	4421	4	US-08-520-373D-6	Sequence 9, Appl1
45	150	6.4	4421	5	PCT-US95-07201-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-08-819-872A-10/c
Sequence 10, Application US/08819872A
Patent No. 6432635
GENERAL INFORMATION:
APPLICANT: Myers, Richard M.
APPLICANT: Cox, David R.
APPLICANT: Pennacchio, Len A.
APPLICANT: Lehesjoki, Anna-Elina
APPLICANT: de la Chappelle, Albert
TITLE OF INVENTION: Cystatin B Mutants
FILE REFERENCE: STAN-104
CURRENT APPLICATION NUMBER: US/08/819,872A
CURRENT FILING DATE: 1997-03-18
PRIOR APPLICATION NUMBER: 60/013,975
PRIOR FILING DATE: 1996-03-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 2822
TYPE: DNA
ORGANISM: Homo sapiens
US-08-819-872A-10

Alignment scores:

Pred. No.: 7.61e-06
Score: 179.00
Percent Similarity: 40.88%
Best Local Similarity: 26.69%
Query Match: 7.62%
DB: 4
Gaps: 12

US-09-622-964-5 (1-435) x US-08-819-872A-10 (1-2822)

140 LeuArgSerValSerThrAlaValTyrGlyArgPheProSerAlaGlnHisLeuValGln 159

```

Db      2111 CTAAGACACACAGCCGGGCGTTCACAGCGGTTTCTTACACAGACCCGTTGGGGCAG 2052
Qy      Ala-----GlyPheMetThrProAlaGlnHisLysGln 170
Db      2051 GCCCTCTTACAGCCGACCTACCTTGTATGAGTAGTTGTCGCCGGACACACCTGGCT 1992
Qy      171 LeuGluLysLeuSerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsn 190
Db      1991 CTTGAATGACAGCGGCTTAACACAGGACCTTGTGTTCTTCTCTCTC-----1941
Qy      191 LeuSerMetLysAlaTrpLeuGlyLysArgile-----ArgAspProIleLeuLeuGln 208
Db      1940 -----AACTGGGACCTCAGCTAGACAGACAGAGAGAAATAGAGATGCT 1896
Qy      209 SerLeuLeuAsnGluMetLysThrLeuArgThrGlnCysGlyHisLysLeuTyraLysAsp 228
Db      1895 CAGTGGCTTCTTGGATTCCTCCCTGCTGCTAGTAACCTGACAGCATGTGAGACACAC-- 1839
Qy      229 TrpIleSerIleProLeuValLysTrpGlnValLysTrpValAlaValLysSerPhePhe 248
Db      1838 -----ACGATCTGACACGCTCTTTCACACAAATTGT- 1807
Qy      249 LeuThrCysLeuValGlyArgGlnPheLeuAsnProAlaLysAlaLysTrpGlyHisGln 268
Db      1806 CTTCTACACACTGAGAGGACAT-----TTAAAC-----TGTACGCTTAATAATAA 1762
Qy      269 LeuAspLeuValValProValPheThrPheLeuGlnPhePhePheTrpValLysTrpLeu 288
Db      1761 GTAGAA-----AAAA--CAGTTCCAGCTGGGTACGGTGT 1730
Qy      289 LysValGlyLeuSerArgAlaLeuLeuGlyTrpArgHisGlyHisLysGlnArgLysGln 308
Db      1729 GCTCAAGATGATTAATCCCACTGAGAGGCTGAGGTGGGAGGCTGAGGAGCTACCTGAGGTGAG 1670
Qy      309 GlnLeuLeuGluThrArgMetGlnCysGlnGluArgLysValSer-----ArgValGlu 326
Db      1669 GAG---TTTGAGACACAGCTAGCCCAATGGCAAAACCCGTTTCTTAAATAATACAAA 1613
Qy      327 SerSerGlnAlaIleTrpTrpArgTrpProValIleProAlaThrArgGlnAlaGlnArgLys 346
Db      1612 ATTAGCTGGGCGGTGGGAGACACCTGCAATCCCAATCCGAGGAGGCTAAGCAGAGA 1553
Qy      347 GluSerLeuGluProGlyLysArgArgLeuTrpTrpGlnSerSerSerSerThrProLeu 366
Db      1552 GAATCCCTTGAACCTGGAGAGGAGGAGGTTGCAGTG-----GCCGAGATTGTGCCACTG 1499
Qy      367 GluArgMetMet-----MetIleLeuArgProThrGly 377
Db      1498 CACTCCACCTGGGCAAGAGTTGAGACTGAGTCTCAAAAACAAAAAACAACACAGGC 1439
Qy      378 LeuSerThrGlyIleCysArgCysProCysTrpLeuTrpMetArgCysThrArgThrCys 397
Db      1438 CTCAGC-----1433
Qy      398 LeuGlyTrpSerArgThrCysThrGlyLysSerProSerHisSerPro 413
Db      1432 ATCGCTGGACCTGTTCTTGTGCTGGGCGATCCCGTGGTCAACACCA 1385

```

RESULT 2

```

US-08-819-872A-1/c
; Sequence 1, Application US/08819872A
; Patent No. 6432635
; GENERAL INFORMATION:
; APPLICANT: MYERS, Richard M.
; APPLICANT: COX, David R.
; APPLICANT: Pennacchio, Len A.
; APPLICANT: Lehesjoki, Anna-Elina
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: Cystatin B Mutants
; FILE REFERENCE: STRAN-104
; CURRENT APPLICATION NUMBER: US/08/819, 872A
; PRIORITY FILING DATE: 1997-03-18
; PRIORITY APPLICATION NUMBER: 60/013, 975

```

```

; PRIOR FILING DATE: 1996-03-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-819-872A-1

Alignment Scores:
Pred. No.: 9,29e-06 Length: 2822
Score: 178.00 Matches: 79
Percent Similarity: 40.88 Conserved: 42
Best Local Similarity: 26.69 Mismatches: 100
Query Match: 7.57 Indels: 77
DB: 4 Gaps: 12

US-09-622-964-5 (1-435) x US-08-819-872A-1 (1-2822)

Qy      140 LeuArgSerValSerThrAlaValLysTrpValArgPheProSerAlaGlnHisLeuValGln 159
Db      2111 CTAAGACACACAGCCGGGCGTTCACAGCGGTTTCTTACACAGACCCGTTGGGGCAG 2052
Qy      160 Ala-----GlyPheMetThrProAlaGlnHisLysGln 170
Db      2051 GCCCTCTTACAGCCGACCTACCTTGTATGAGTAGTTGTCGCCGGACACACCTGGCT 1992
Qy      171 LeuGluLysLeuSerLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsn 190
Db      1991 CTTGAATGACAGCGGCTTAACACAGGACCTTGTGTTCTTCTCTCTC-----1941
Qy      191 LeuSerMetLysAlaTrpLeuGlyLysArgile-----ArgAspProIleLeuLeuGln 208
Db      1940 -----AACTGGGACCTCAGCTAGACAGACAGAGAGAAATAGAGATGCT 1896
Qy      209 SerLeuLeuAsnGluMetLysThrLeuArgThrGlnCysGlyHisLysLeuTyraLysAsp 228
Db      1895 CAGTGGCTTCTTGGATTCCTCCCTGCTGCTAGTAACCTGACAGCATGTGAGACACAC-- 1839
Qy      229 TrpIleSerIleProLeuValLysTrpGlnValLysTrpValAlaValLysSerPhePhe 248
Db      1838 -----ACGATCTGACACGCTCTTTCACACAAATTGT- 1807
Qy      249 LeuThrCysLeuValGlyArgGlnPheLeuAsnProAlaLysAlaLysTrpGlyHisGln 268
Db      1806 CTTCTACACACTGAGAGGACAT-----TTAAAC-----TGTACGCTTAATAATAA 1762
Qy      269 LeuAspLeuValValProValPheThrPheLeuGlnPhePhePheTrpValLysTrpLeu 288
Db      1761 GTAGAA-----AAAA--CAGTTCCAGCTGGGTACGGTGT 1730
Qy      289 LysValGlyLeuSerArgAlaLeuLeuGlyTrpArgHisGlyHisLysGlnArgLysGln 308
Db      1729 GCTCAAGATGATTAATCCCACTGAGAGGCTGAGGTGGGAGGCTGAGGAGCTACCTGAGGTGAG 1670
Qy      309 GlnLeuLeuGluThrArgMetGlnCysGlnGluArgLysValSer-----ArgValGlu 326
Db      1669 GAG---TTTGAGACACAGCTAGCCCAATGGCAAAACCCGTTTCTTAAATAATACAAA 1613
Qy      327 SerSerGlnAlaIleTrpTrpArgTrpProValIleProAlaThrArgGlnAlaGlnArgLys 346
Db      1612 ATTAGCTGGGCGGTGGGAGACACCTGCAATCCCAATCCGAGGAGGCTAAGCAGAGA 1553
Qy      347 GluSerLeuGluProGlyLysArgArgLeuTrpTrpGlnSerSerSerSerThrProLeu 366
Db      1552 GAATCCCTTGAACCTGGAGAGGAGGAGGTTGCAGTG-----GCCGAGATTGTGCCACTG 1499
Qy      367 GluArgMetMet-----MetIleLeuArgProThrGly 377
Db      1498 CACTCCACCTGGGCAAGAGTTGAGACTGAGTCTCAAAAACAAAAAACAACACAGGC 1439
Qy      378 LeuSerThrGlyIleCysArgCysProCysTrpLeuTrpMetArgCysThrArgThrCys 397

```


FEATURE:
NAME/KEY: exon 1
LOCATION: 990..1372
FEATURE:
NAME/KEY: exon 2
LOCATION: 2115..2262
FEATURE:
NAME/KEY: exon 3
LOCATION: 3056..3202
FEATURE:
NAME/KEY: exon 4
LOCATION: 3387..3472
FEATURE:
NAME/KEY: exon 5
LOCATION: 5667..5923
FEATURE:
NAME/KEY: exon 6
LOCATION: 7745..8955
US-09-077-354B-3

Alignment Scores:
Pred. No.: 0.000743 Length: 10380
Score: 166.00 Matches: 57
Percent Similarity: 38.50% Conservative: 15
Best Local Similarity: 30.48% Mismatches: 48
Query Match: 7.06% Indels: 67
Gaps: 9

US-09-622-964-5 (1-435) x US-09-077-354B-3 (1-10380)
QY 301 H1SGLYGlnArgGlyHisGlyGlnGlnLeu---LeuGluThrArgMetGlnGlyGlnGlu 319
Db 6833 CACTCCAGCCTGGGTGACACAGCGAGAGCTTCTCAAAACAAACAAATTGC----- 6780
QY 320 ArgLysValSerArgValGlnSerSerGlnAlaTyrPheArgPheProValIleProAla 339
Db 6779 -----CGGGCGTGGTGGCGGCTGCTAAGTCCACCT 6747
QY 340 ThrArgGlnAlaGlnAlaGlyLysSerLeuGluProGlyArgArgLeuTyrPheGln 359
Db 6746 ACTCGGAGGCTGAGGACAGAGAAATCGCTTGAACCTAGGACAGCAAGTTG---CAG 6693
QY 360 SerSerSerSerThrProLeu----- 366
Db 6692 TGAGATGAGATGTGCCACTGTGCTCCAGCCTCAGCAACAGAGTGAATCTGTCTCAAA 6633
QY 367 ---GluArgMetMetMetIleLeu-----Arg-ProThrGlyLeu----- 378
Db 6632 AAAAAAAATGTAGCTTCTCTCAGGCTTGGGTAGGGCTGAGAAATTGCCATCTCA 6573
QY 379 -SerThrGlyIleCysArgCysProCysTyrLeuTyrMetArgCysThrArgThrCysLe 398
Db 6572 CAGCTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6513
QY 398 uGlyTyrSerArgThrCysThrGlyIleSerProSerHisSerProProThrGlnLeu-- 417
Db 6512 C-----AAGAAAGCTGAGCTGATCTTCCCTTCTGGGCAACAGAGCGCTCCCA 6459
QY 418 -----LeuProProSerSerValIleProPr 426
Db 6458 CTAGTCATTCCTTACTTGGGCGGCTCTCTCCTCCTCCTCCTTAAAGTGGCAGAGA 6399
QY 426 OLeuTyr----- 429
Db 6398 GCCCTGGCTTCACAGTTTGTACAGAGATCGCAGCTTAGTGCCTTAAAGCAGATTCACATC 6339
QY 429 aProProSerThrSerAla 435
Db 6338 ACCTCGGCTGCGAGTCC 6320

RESULT 5
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A

Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Krommal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A. 35,136
REGISTRATION NUMBER: 017957-000100
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..246240
OTHER INFORMATION: /note="HLA-H.CONTIG"
US-08-724-394A-20
Alignment Scores:
Pred. No.: 0.0945 Length: 246240
Score: 166.00 Matches: 74
Percent Similarity: 35.76% Conservative: 29
Best Local Similarity: 25.69% Mismatches: 83
Query Match: 7.06% Indels: 103
Gaps: 13
US-09-622-964-5 (1-435) x US-08-724-394A-20 (1-246240)
QY 229 TrpIleSerIleProLeuValIleTyrThrGlnValIleThrValAlaValIleSerPhePhe 248
Db 39543 TGGGAAGCCTTCCC---TGATACCCCGAGTGGCAGAGGCTTCATGTCTTTCTTA 39487
QY 249 LeuThrCysLeuValIleGlyArgGlnPheLeu----- 258
Db 39486 GTACACCTGCAAGTTT-GTATTGTTCATGTTCATTCCTTCACTTCTTACTTGTCTCT 39428
QY 259 -----AsnProAlaLysAlaTyrProGlyHisGlu 268
Db 39427 CTGTGTGCTCTCATTCAGTGCAGCTGAACTTATGAAGCATCTCATGGTGCAGAT 39368
QY 269 LeuAspLeuValIle-----ProValPheThrPheLeuGlnPhePhePhe-- 283
Db 39367 CTTAATAATTAAATATGTGCGAAGCTAATGTCTATGATACAGAAATTTATCA 39308

```

QY 284 -----TyrValGlyTyrLeuLys-----ValGlyLeuSer 293
Db 39307 AAAAAAATATAGTATGTTGGTGGCCAGTGAACAGCCCGTAATCCAGCACTTGG 39248
QY 294 ArgAlaLeuLeuGlyTyrPargHisGlyGlnArgGlyHisGlyGln----- 308
Db 39247 GAGCCGAGGAGAGAGATCATAGAGTCAAGAAATTCAGAGCCAGCCGCAAAATGG 39188
QY 309 ---GlnLeuLeuGlnThrArgMetGlnCysGlnGlnArgLysValSerArgValGluSer 327
Db 39187 TGAACCTCATCTCTACTATA-----AATACAAAAAGT 39155
QY 328 SerGlnAlaLeuThrPargThrPargProValIleProAlaThrArgGlnAlaGlnAlaGlyGlu 347
Db 39154 AGCCAGCGGTGGTGGTGGCCAGCTGTAATCCAGTACTCAGAGAGGCTGAAGCGGAGGA 39095
QY 348 SerLeuGlnProGlyArgArgArgLeuTyr-----TyrGlnSerSer 361
Db 39094 TCACCTGAACTGGAGGAGAGATTCAGATGAGTGAATGATGCTGACCTGACCTCCAGC 39035
QY 362 -----SerSerThrProLeuGln----- 367
Db 39034 CTGGGCGACAGTGAAGTCTCATCTCAAAATATATATATATATATATATATATAT 38975
QY 368 -----ArgMetMetMet 371
Db 38974 AATAATTTATGAAATGAACTGCTGATTTGAAATAGCTGTTTTTAAAAATTAATA 38915
QY 372 IleLeuArgProThrGlyLeuSerThrGly-----IleCysArgGlyProCys 387
Db 38914 TTTTATTAAGTCTGAGTACAGTACAGAGATGTGAGGTTTCTTACATAGGTAACGTGT 38855
QY 388 TrpLeuTyrMetArgCysThr-----ArgThrCysLeu 398
Db 38854 GCCATGGTATTTGCTGCACTATCAACCACTACCTAGTATTAAGTACAGCATGCTT 38795
QY 399 GlyTyrSerArgThrCysThrGlyIleSerProSerHisSerProThrGlnLeuLeu 418
Db 38794 AGCTCTTTTACCTAATGTTTCCACACACCCCACTCTCCCAACAGGC----- 38741
QY 419 ProProSerSerValGluProPro 426
Db 38740 CCCAGTGAAGTGTGTTCCCTCC 38717

RESULT 6
US-08-724-394A-21/C
Sequence 21, Application US/08724394A
Patent No. 5872237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kironmal, Gregory S.
APPLICANT: Laufer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies thereo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Flets, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H. CONTIG"

US-08-724-394A-21

Alignment Scores:
Pred. No.: 0.0945
Score: 166.00
Percent Similarity: 35.76%
Best Local Similarity: 25.69%
Query Match: 7.06%
Gaps: 13

US-09-622-964-5 (1-435) x US-08-724-394A-21 (1-246240)

QY 229 TrpIleSerIleProLeuValTyrThrGlnValAlaValAlaValTyrSerPhePhe 248
Db 39543 TGGGAAAGCTTCCCTCCCTGATACCCCACTTGGCAGAGATCTTCACTTTCTTTCTA 39487
QY 249 LeuThrCysLeuValGlyArgGlnPheLeu----- 258
Db 39486 GTCACTGTGCAAGTTT-GTATTTGTCATGTTTATCATCTTCATCTGCTGCTC 39428
QY 259 -----AsnProAlaValAlaTyrProGlyHisGlu 268
Db 39427 CTGTGTGTGTCATTCAGTGCAGCTGCACTTATGAAAGCATGATGCTGCTCAGAT 39368
QY 269 LeuAspLeuValVal-----ProValPheThrPheLeuGlnPhePhePhe 283
Db 39367 CTTAATTAATTAATATTGTCGAAAGTATGTCATGTCATGATGATGATGATGATG 39308
QY 284 -----TyrValGlyTyrLeuLys-----ValGlyLeuSer 293
Db 39307 AAAAAAATATAGTATGTTGGTGGCCAGTGAACAGCCCGTAATCCAGCACTTGG 39248
QY 294 ArgAlaLeuLeuGlyTyrPargHisGlyGlnArgGlyHisGlyGln----- 308
Db 39247 GAGCCGAGGAGAGAGATCATAGAGTCAAGAAATTCAGAGCCAGCCGCAAAATGG 39188
QY 309 ---GlnLeuLeuGlnThrArgMetGlnCysGlnGlnArgLysValSerArgValGluSer 327
Db 39187 TGAACCTCATCTCTACTATA-----AATACAAAAAGT 39155
QY 328 SerGlnAlaLeuThrPargThrPargProValIleProAlaThrArgGlnAlaGlnAlaGlyGlu 347
Db 39154 AGCCAGCGGTGGTGGTGGCCAGCTGTAATCCAGTACTCAGAGAGGCTGAAGCGGAGGA 39095
QY 348 SerLeuGlnProGlyArgArgArgLeuTyr-----TyrGlnSerSer 361
Db 39094 TCACCTGAACTGGAGGAGAGATTCAGATGAGTGAATGATGCTGACCTGACCTCCAGC 39035
QY 362 -----SerSerThrProLeuGln----- 367
Db 39034 CTGGGCGACAGTGAAGTCTCATCTCAAAATATATATATATATATATATATATATAT 38975

```

```

QY 368 -----ArgMetMetMet 371
Db 38974 AATAATGTGATGGAATGAACTGCTGATGGAATAAGCTGTTTAAATAATATTA 38915
QY 372 ILeuAArgProThrGlyLeuSerThrGly-----IleCysArgCysProCys 387
Db 38914 TTTTAAAGTCTCGGTACAGACAGATGTCAGGTTGTTACATAGGTAAACGTGT 38855
QY 388 TrpLeuTrpMetArgCysThr-----ArgThrCysLeu 398
Db 38854 GCCATGGTGAATGCTGCTACCTATCAACCATCATCAGTATTAAGTACAGATGAT 38795
QY 399 GlyTrpSerArgThrCysThrGlyIleSerProSerHisSerProProThrGlnLeu 418
Db 38794 AGCTCTTTACCTAATGTTCTCCACACCCCACTCCCTCCCAACAGCC----- 38741
QY 419 ProProSerSerValGluProPro 426
Db 38740 CCCAGTGAAGTGTGTTCCCTCC 38717

RESULT 7
US-08-724-394A-22/c
; Sequence 22, Application US/08724394A
; Patent No. 587237
GENERAL INFORMATION:
APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolf, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
TITLE OF INVENTION: Sequences and Antibodies Thereto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESS: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724.394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONFIC"
US-08-724-394A-22
Alignment Scores: 0.0945 Length: 246240
Pred. No.:

```

```

Score: 166.00 Matches: 74
Percent Similarity: 35.76% Conservative: 29
Best Local Similarity: 25.69% Mismatches: 83
Query Match: 7.06% Indels: 103
DB: 2 Gaps: 13

US-09-622-964-5 (1-435) x US-08-724-394A-22 (1-246240)
QY 229 TrpIleSerIleProLeuValTrpGlnValAlaValAlaValTyrSerPhe 248
Db 39543 TGGGAAGCCCTCC-----TGATACCCCAAGTGGCAGAGAGCTTCATTTGCTTTCTA 39487
QY 249 LeuThrCysLeuValGlyArgGlnPheLeu----- 258
Db 39486 GTCACCTGTGCAAGTTT-GTATGTCATGTTTATCATCCCTCATTCAGTGTCTGTCT 39428
QY 259 -----AsnProAlaLysAlaTyrProGlyHisGlu 268
Db 39427 CTGTGTGTCCTCATTCAGTGAAGTCTGAACTGTTTGAAGTCAATGTCATGTCAGAT 39368
QY 269 LeuAspLeuValAla-----ProValPheThrPheLeuGlnPhePhe--- 283
Db 39367 CTATATTAATTAATTAATGTCGGAAGCTAATGTCATGTCAGATTAAGAAATTAATCA 39308
QY 284 -----TyrValGlyTrpLeuLys-----ValGlyLeuSer 293
Db 39307 AAAAAAATATAGTATGTTGGTGGCCAGTCAAGCCGTAATCCAGACCTTTGG 39248
QY 294 ArgAlaLeuLeuGlyTrpArgHisGlyGlnArgGlyHisGlyGln----- 308
Db 39247 GAGCCGAGGAGGAGGATCATGAGAGTGAATTAATCAAGACAGCTGGCCAAATGG 39188
QY 309 --GlnLeuLeuGlnTrpArgMetGlnCysGlnGlnArgLysValSerArgValGluSer 327
Db 39187 TGAACCTCATCTCTATTA-----AATACAAAAAGT 39155
QY 328 SerGlnAlaTrpTrpArgTrpProValIleProAlaThrArgGlnAlaGlnArgLys 347
Db 39154 AGCCAGGCGTGGTGGTCCCACTGTATCCAGTCTCAGAGAGGTGAAGGGAGCA 39095
QY 348 SerLeuGlnProGlyArgArgArgLeuTrp-----TrpGlnSerSer 361
Db 39094 TCACCTTAACCTGGAGGAGAGATGCAATGCAATGAGATCATGCCACTGCACCTCCAGC 39035
QY 362 -----SerSerThrProLeuGlu----- 367
Db 39034 CTGGGCGACAGTGAAGCTCATCTCAAAATTAATTAATTAATTAATTAATTAAT 38975
QY 368 -----ArgMetMetMet 371
Db 38974 AATAATGTGATGGAATGAACTGCTGATGGAATAAGCTGTTTAAATAATATTA 38915
QY 372 ILeuAArgProThrGlyLeuSerThrGly-----IleCysArgCysProCys 387
Db 38914 TTTTAAAGTCTCGGTACAGACAGATGTCAGGTTGTTACATAGGTAAACGTGT 38855
QY 388 TrpLeuTrpMetArgCysThr-----ArgThrCysLeu 398
Db 38854 GCCATGGTGAATGCTGCTACCTATCAACCATCATCAGTATTAAGTACAGATGAT 38795
QY 399 GlyTrpSerArgThrCysThrGlyIleSerProSerHisSerProProThrGlnLeu 418
Db 38794 AGCTCTTTACCTAATGTTCTCCACACCCCACTCCCTCCCAACAGCC----- 38741
QY 419 ProProSerSerValGluProPro 426
Db 38740 CCCAGTGAAGTGTGTTCCCTCC 38717

RESULT 8
US-09-050-159-129/c
; Sequence 129, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:

```


Patent No. 5840506
GENERAL INFORMATION:
APPLICANT: Giordano, Antonio
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS AND PROGNOSIS OF
NUMBER OF INVENTIONS: CANCER
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEIDEL, GONDA, LAVORGNA & MONACO, P.C.
STREET: Suite 1800 Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/832,877
APPLICATION NUMBER: US/08/832,877
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 8321-13 US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-5549
TELEFAX: (215) 568-8383
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 3507 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-832-877-67

Alignment Scores:
Pred. No.: 0.00421
Score: 160.50
Percent Similarity: 62.67%
Best Local Similarity: 48.00%
Query Match: 6.83%
DB: 2
Gaps: 3

US-09-622-964-5 (1-435) x US-08-832-877-67 (1-3507)

QY 300 ArgH1sglyglnArgGlyH1sglyglnGlnLeuGlnThArgMetGlnCysGlnGlu 319
DB 2886 CAGCATGGCCACATGCTAAACCCGATCTCTACTA----- 2851
QY 320 ArgGlyValSerArgValGluSerSerGlnAlaTrpTrpArgThrProValIleProAla 339
DB 2850 ---AAAATAACAAAATTT-----AGCTGGGGCTGTGGCGGACCTGTATCCCACT 2800
QY 340 ThArgGlnAlaGlnAlaGlnGluSerLeuGlnProGlyArgArgGlyLeuTrpGln 359
DB 2799 ACTGAGGAGCTGTGGCGGAGAAATCACTTAACCTGGAGCGGAGAGTTG-----CAG 2746
QY 360 SerSerSerSerTrpProLeuGlnArgMetMetIleLeuArg 374
DB 2745 TGAGCCGAGATCACGCCACTGCTCACTCCAGCTGGCGCAATGCA 2701

RESULT 13
US-09-128-155-17
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: 09404/052001

CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Alignment Scores:
Pred. No.: 0.279
Score: 158.00
Percent Similarity: 37.31%
Best Local Similarity: 28.36%
Query Match: 6.72%
DB: 3
Gaps: 7

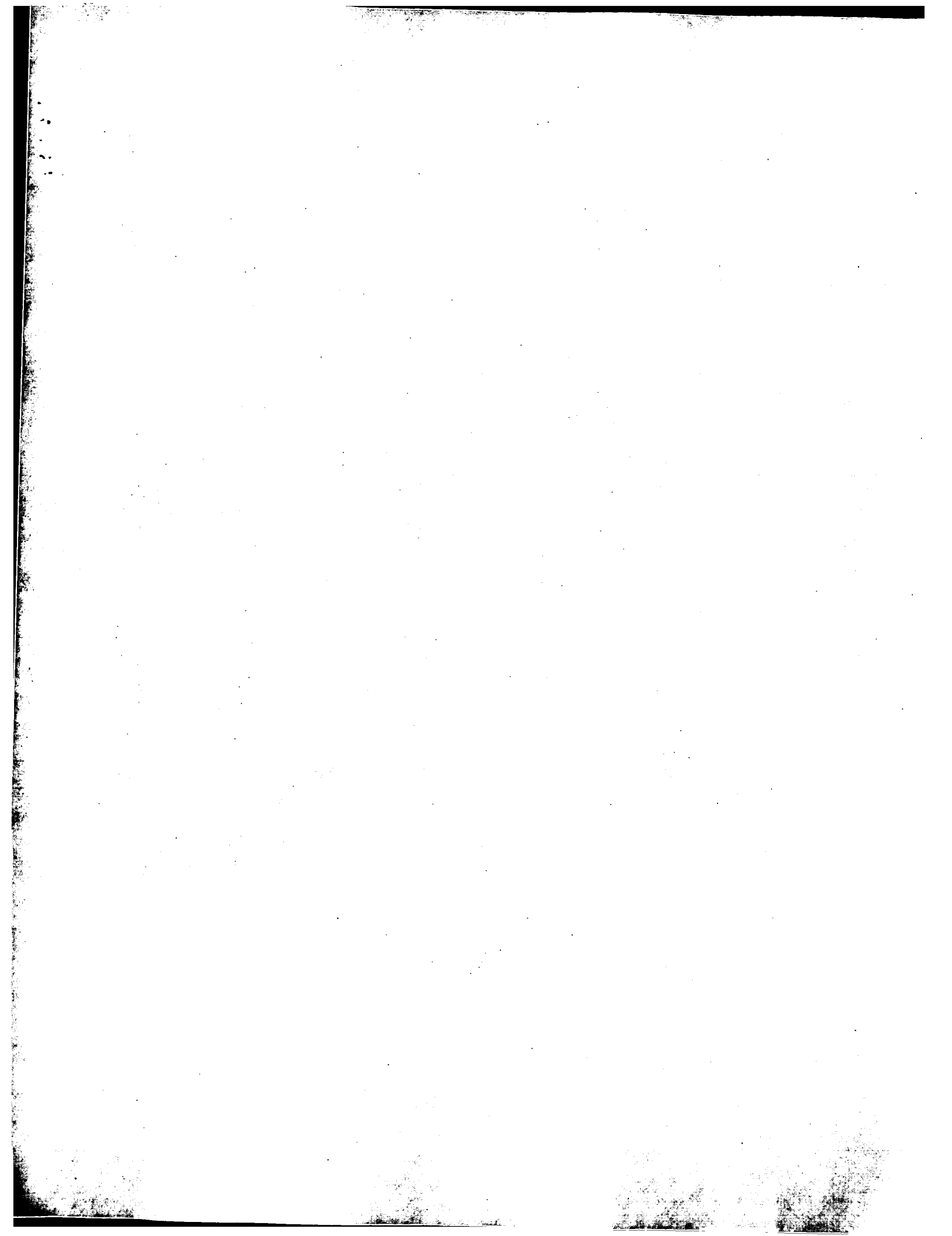
US-09-622-964-5 (1-435) x US-09-128-155-17 (1-176373)

QY 299 TrpArgH1sglygln-----ArgGlyH1sglyglnGln 309
DB 170504 TTGATAGGGGCTCAAGTAATTTCTACTTAAATAATACAAAGGGCTGGCGGTGC 170563
QY 310 LeuLeuGlnThr-----ArgMetGlnCysGlnGlnAlaArgValSerArg 324
DB 170564 CTCATGCTCTATATCCCGCATTTGGAGGCTGAGAGCGCGATCAAGGTCAAGAGA 170623
QY 325 Val-----GluSerSer 328
DB 170624 TTGAGACCATCCTGGCCAAACGCGTAACCGTCTCTCACTTAAATAATACAAATTAGC 170683
QY 329 GlnAlaTrpTrpArgThrProValIleProAlaThrArgGlnAlaGlnGlyGluSer 348
DB 170684 TGGCATGTGTGTCGTCCTTATTCCTCCACTGCGGAGGCTGAGCGAGAGATCG 170743
QY 349 LeuGlnProGlyArgArgArgLeuTrpTrpGlnSer-----Ser 361
DB 170744 CTTGACCTGGAGGTGAGGTTTCACAGTGGCGGAGATCGCACCTGCAATCCAGAGCG 170803
QY 362 -----SerSer 364
DB 170804 AGACTGTGTCAAAAAAAAAAAAAAAAAAATAATCCAAACAGATGACCTTAAGCTGCA 170863
QY 364 TrpLeuGln-----ArgMetMetIleLeuArgProThrGlyLeuSerThrGly 382
DB 170864 GAGCTTGAGACATCTAGGAGCTGATGATCAGTACACAAACATTAATGTGTCACAGGCT 170923
QY 382 eCyArgCysProCysTrpLeuTrpMetArgCysThrArgThrCysLeuGlyTrpSer 402
DB 170924 GATGAATGCACAGCAGACAGACTTCAGATGATGATCAGTCAAGTATA----- 170969
QY 402 gThrCysThrGlyIleSerProSerHisSerProProThrGlnLeu----- 418
DB 170970 -----TCCACAAGTCCACCTCAAGAAATGCTATATCAACAT 171007
QY 419 -----ProProSerSerValGluProProLeuTrpAlaProProSerThr 434
DB 171008 TTGGCATCAATCTCTATCAAAAGATAGTCCAAAGCATGGTTCACAAACACTTTC 171067
QY 434 r 434
DB 171068 c 171068

RESULT 14


```
Db 2617 ACAC----- 2620
QY 175 rLeuProHisAsnMetPheTrpValProTrpValTrpPheAlaAsnLeu----- 191
Db 2621 ---AGTGAGCATGCGCTGTGTC-CCATCTACTTGGAGGCTAAGGTGAGAGGTCCT 2675
QY 192 -----SerMetLysAlaTrpLeuG1 198
Db 2676 TGAGCCCGAGAGCTTAAGCTGCAGTAGCCATGATCTCACCATTTGGCATTCCTCGG 2735
QY 198 YG1AArg1LeuArgAspPro1LeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuAr 218
Db 2736 CAACAGAGCTAGA---CCCTGTCTCTTAAAAAATAATA----- 2772
QY 218 gThnGlnCysGlyHisLeuTrpAlaTrpAspTrp1LeSer1LeProLeuValTrpThrG1 238
Db 2773 -----ATATATATATACGA 2786
QY 238 nValValThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLe 258
Db 2787 AATAATAGTAGAATAATGTATGATGTTTATCAGAAATCTGTATCCTTACTTATTT 2846
QY 258 uAsnProAlaLysAlaTyrProG1HisGluLeuAspLeuValAlaProValPheThrPh 278
Db 2847 A-----TATCCATGATGAAATAATTAAAAACAGACCCAAACAGGTGAT 2891
QY 278 eLeuGlnPhePhePheTyrValGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeuG1 298
Db 2892 TTTAAAACTTAACCTTGTGCGGCTGTGCTCACACCTGTAACTTCAGCAGCTTGGG 2951
QY 298 yTrpArgHisGlyGlnArgGlyHisGlyGlnGlnLeuLeuGluThrArgMet-----G1 316
Db 2952 AGGCGGAGGTGGCGAGATCACCTGAGTGCAGAG---TTCAAGACACAGCTGGCCACAT 3008
QY 316 nCysGlnGluArgLysValSerArgValGluSerSerGlnAlaTrpTrpArgThrProVa 336
Db 3009 GCGGAACCCGCTCT-CTACTAAATAATAATAATTAACGTGGGTGTGTCAGCGCTGT 3067
QY 336 11LeProAlaThrArgGluAlaGluAlaGlyGluSerLeuGlnProGlyArgArgArgLe 356
Db 3068 AATCCACAGCTACTGAGAGGCTGAGGCGAGAGATCGCTTAACCTGGGAAGCAGAGGTT 3127
QY 356 uTrp-----TrpGlnSerSerSerSerThrProLe 366
Db 3128 GCACTGAGCCAGATCGCCCATTCGACTTACGCTGGCGAAACAAGAGTGAAACTCCGTC 3187
QY 366 uGluArgMetMetLeuLeu 373
Db 3188 TCAAAAGAAAAAATTACTT 3209
```

Search completed: July 27, 2003, 14:57:24
Job time : 434.717 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

SUMMARIES

OM protein - nucleic search, using frame-plus-p2n model
Run on: July 27, 2003, 11:06:55 ; Search time 261.665 Seconds
(without alignments)
3743.798 Million cell updates/sec

Title: US-09-622-964-5
Perfect score: 2350
Sequence: 1 MTITYSQVANAARLGSFSL.....DLPPSYVEPLMAPPTSA 435

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seg length: 0
Maximum DB seg length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame-p2n.model -DEV-xlh
-Q/cgn2_1/USPTO.spool/US0962264.runtat.22072003.101118.26006/app.query.fasta.1.2069
-DB-N-Geneseq.101002 -OPMT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOCLIGN=200 -THR.SCORE=100 -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO-MAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV-TIMEOUT=120 -WARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N-Geneseq.101002.*
1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	2350	100.0	2429	20	AAZ21228	Human CGICE long f
2	1925	81.9	2229	20	AAZ21227	Human CGICE short
3	1488	63.3	1916	20	AAZ21229	Mouse CGICE cDNA s
4	1302	55.4	1837	22	ABAI4558	Human nervous syst
5	1293	55.0	1650	22	ABAI4559	Human nervous syst
6	1293	55.0	18530	22	ABAI4557	Human nervous syst
7	1282	54.6	16125	20	AAZ21226	Human CGICE genom
8	1216	51.7	7108	22	ABAI4556	Human nervous syst
9	801	34.1	2861	23	ABLI10793	Drosophila melanog
10	685	29.1	1608	23	ABLI10792	Drosophila melanog
11	661	28.1	10760	23	ABLI10791	Drosophila melanog
12	644	27.4	3592	23	ABLI12636	Drosophila melanog
13	600	25.5	3785	23	ABLI12637	Drosophila melanog
14	560.5	23.9	1345	23	ABLI12638	Drosophila melanog
15	514	21.9	2196	23	ABLI16195	Drosophila melanog
16	495	21.1	1526	23	AAZ21236	Drosophila melanog
17	482	20.5	7505	23	AAZ21237	Drosophila melanog
18	462	19.7	1350	22	AAZ21238	Drosophila melanog
19	448	19.1	305	22	ABAI6848	Human secreted pro
20	432	18.4	1263	20	AAZ21239	Human nervous syst
21	339.5	14.4	2104	23	AAZ21240	Human adult retina
22	270	11.5	392	22	AAZ21241	DNA encoding novel
23	270	11.5	401	22	AAZ21242	Human breast cancer
24	263	11.2	527	22	AAZ21243	Human breast cancer
25	238	10.1	507	22	AAZ21244	Human breast cancer
26	238	10.1	507	22	AAZ21245	Human breast cancer
27	220	9.4	1662	23	AAZ21246	Human digestive sy
28	194.5	8.3	13743	22	AAZ21247	DNA encoding novel
29	189	8.0	10820	22	ABAI6598	Genomic sequence f
30	183.5	7.8	8056	22	AAZ21248	Human nervous syst
31	183	7.8	10302	22	AAZ21249	Human nervous syst
32	182	7.7	13670	24	ABAI6133	Human nervous syst
33	182	7.7	13670	24	ABAI6133	Human nervous syst
34	179.5	7.6	12822	22	AAZ21250	Human nervous syst
35	179	7.6	107820	22	AAZ21251	Human nervous syst
36	178.5	7.6	9457	24	ABAI6539	Human nervous syst
37	178	7.6	2821	18	AAZ21252	Human nervous syst
38	178	7.6	2821	18	AAZ21253	Human nervous syst
39	178	7.6	2821	18	AAZ21254	Human nervous syst
40	178	7.6	2821	18	AAZ21255	Human nervous syst
41	178	7.6	2821	18	AAZ21256	Human nervous syst
42	177.5	7.6	19199	22	AAZ21257	Human nervous syst
43	176	7.5	17047	22	AAZ21258	Human nervous syst
44	175.5	7.5	145831	24	ABAI62309	Human nervous syst
45	175.5	7.5	145831	24	ABAI62309	Human nervous syst

ALIGNMENTS

RESULT 1
AAZ21228
AAZ21228 standard: cDNA: 2429 BP.
AC
AAZ21228;
XX
22-NOV-1999 (first entry)
DT
XX
Human CGICE long form cDNA sequence.
DE
XX
CGICE: Best's macular dystrophy; mutation; diagnosis; detection;
KW
BMD; age-related macular dystrophy; ss.
KW
XX
Homo sapiens.
OS
XX
Key
CDS
Location/Qualifiers
105..1412
/*tag= a

/Product= "CGICE short form protein"

W09943695-A1.

02-SEP-1999.

22-FEB-1999; 99WO-US03790.

25-FEB-1998; 98US-0075941.

18-DEC-1998; 98US-0112926.

(MERI) MERCK & CO INC.

(UYUP-) UNIV UPPSALA.

Petrushin K, Caskey CT, Metzker M, Madellius C;

WPI: 1999-540560/45.

P-PSDB: AAY29954.

Human and mouse polynucleotides encoding CGICE polypeptides -

Claim 2; Fig 4; 67pp; English.

The present sequence represents the human CGICE cDNA sequence, which when mutated is responsible for Best's macular dystrophy (BMD). CC polynucleotides encoding CGICE are useful for diagnosing whether a CC patient carries a mutation in the CGICE gene. Normal and mutated CC CGICE proteins are useful for identifying activators and/or inhibitors of these proteins, in order to treat BMD. The CGICE gene offers a CC simpler and cheaper method of diagnosing BMD without the need for the CC presence of the patient. The gene may also be useful to discovering the genetic cause of age-related macular dystrophy.

SQ Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 other;

Alignment Scores:

Pred. No.:

5, 51e-163

Score: 2350.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 20

US-09-622-964-5 (1-435) x AA221228 (1-2429)

QY 1 MetThrIleThrThyThSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 DB 105 ATGACCATCACTTACACAAAGCCAGTGGCTAATGCCGCTTAGGCTCTTCCCGCCTG 164
 QY 21 LeuLeuCystrParGlySerIleThyLysLeuLeuThyGlyGlnPheLeu 40
 DB 165 CTGCTGCTGGCGGGGAGCATCTACAAAGCTGCTAATAGCGAGTCTTATATCTTCCTG 224
 QY 41 LeuGlyThyIleIleIleArgPheIleThyArgLeuAlaLeuThyGlnGlnGlnLeu 60
 DB 225 CTGCTCTACATCATCCGCTTATATAGCGTGGCCCTCAGGAGAAACACAGCTG 284
 QY 61 MetPheGlyLysLeuThyThyLysCysAspSerThyIleGlnLeuIleProIleSerPhe 80
 DB 285 ATGTTTGAAACAGTACTCTTATTTGAGCAGCTACATCCAGCTCATCCCATTTCCCTTC 344
 QY 81 ValLeuGlyPheThyValThrLeuValValThrArgTrpTrpPheGlnGlnLeu 100
 DB 345 GTGCTGGCTCTCTACGTAGCGCTGCTGTGACCGCGTGGTGAACCAAGTACGAACTG 404
 QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGln 120
 DB 405 CCGTGGCCGACCCGCTCATGAGCGCTGTGCGGCTTCCTCGAAGGCAAGACGACAA 464
 QY 121 GlyArgLeuLeuArgArgThyLeuIleArgThyAlaAsnLeuGlnValLeuIleLeu 140
 DB 465 GGGCGGCTGCGGGCGGCGACGCTATCCGCTACGCCCACTGGGCAACGTGCTCATCTG 524

QY 141 ArgSerValSerThrAlaValThyLysArgPheProSerAlaGlnIleLeuValGlnAla 160
 DB 525 CGCAGGCTCAGCACCCGAGCTACAAAGCGCTTCCACAGCGCCACAGCTGTGGCAAGCA 584
 QY 161 GlyPheMetThrProAlaGlnIleLysGlnLeuGlyLysLeuSerLeuProIleAsnMet 180
 DB 585 GGCCTTATGACTCCGGCAGAACACAGAGTGTGAAATGAGCTTACCACACACATG 644
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
 DB 645 TTCTGGGTCCCGGTGTGTGTGGTTCGCAACTCTCATATGAAAGCGGTGGAGGTGCA 704
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAspThrLeuArgTrpGln 220
 DB 705 ATCCGGGACCCATCTCTGCTCCAGACCTCTGCAAGAGATGAAACCTGTGGTACTGAG 764
 QY 221 CysGlyHisLeuThyAlaThyAspTrpIleSerIleProLeuValThyTrpGlnVal 240
 DB 765 TGTGGACACCTGATGCTACAGCATGATGATGATGATGATGATGATGATGATGATGATG 824
 QY 241 ThrValAlaValThySerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 DB 825 ACTGTGGCGGTACAGCTTCTCTGACTGTGTAGTTGGCGGACGTTCTGAACCA 884
 QY 261 AlaLysAlaThyProGlyHisGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
 DB 885 GCCAAGGCTACCTGGCGCATAGGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
 QY 281 PhePhePheThyValGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeuGlyTrpArg 300
 DB 945 TTCTCTCTCTAGT 1004
 QY 301 HisGlyGlnArgGlyHisGlyGlnGlnLeuLeuGlnTrpArgMetGlnCysGlnGlnArg 320
 DB 1005 CATGGCCAGAGGGGTGATGAGGCGCCAGCTGCTTACAGAGAGATGCTCAGAGAAAG 1064
 QY 321 LysValSerArgValGlnSerSerGlnAlaTrpTrpArgTrpProValIleProAlaThr 340
 DB 1065 AAGGCTCAGCGGTGAAGAACAGCAGCGGTGTGGGCGCACACCTGTAATCCAGTACT 1124
 QY 341 ArgGlnAlaGlnAlaGlyGlnSerLeuGlnProGlyArgArgArgLeuTrpGlnSer 360
 DB 1125 CGGAGGCTGAGGAGAGAGATGCTTGAACCCGGAGGCGGAGGTGTGTGTGTGTGTGT 1184
 QY 361 SerSerSerThrProLeuGlnArgMetMetIleLeuArgProThrGlyLeuSerThr 380
 DB 1185 AGCTCATCAACCCCTTTGGAGAGATGATGATGATGATGATGATGATGATGATGATG 1244
 QY 381 GlyIleCysArgCysProCysTrpLeuTrpMetArgCysThrArgTrpCysLeuGlyTrp 400
 DB 1245 GGAATTTGACAGGTGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1304
 QY 401 SerArgThrCysThrGlyLysSerProSerHisSerProThrGlnLeuLeuProPro 420
 DB 1305 AGCGGAGCATAGTCTGGAATTAAGCCGAGACACAGCCCTTCACAGCTGCTTCGCC 1364
 QY 421 SerSerValGlnProProLeuTrpAlaProProSerThrSerAla 435
 DB 1365 AGTTCCGTCAGAGCTCTTATAGGGCTCCACCTTCAACATCAACCC 1409

RESULT 2
 AA221227 standard; cDNA; 2229 BP.
 AA221227;
 22-NOV-1999 (first entry)
 Human CGICE short form cDNA sequence.
 DE CGICE; Best's macular dystrophy; mutation; diagnosis; detection;
 KW BMD; age-related macular dystrophy; ss.

```

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 105..1862
FT /tag=A
FT /product="CGICE long form protein"
FT /trans_except=(pos:465..467,aa:Ser)
XX MO9943695-A1.
XX 02-SEP-1999.
XX 22-FEB-1999; 99MO-US03790.
XX 25-FEB-1998; 98US-0075941.
XX 18-DEC-1998; 98US-0112926.
XX (MERI) MERCK & CO INC.
XX (UYUP-) UNIV OFPSALA.
XX Petrukhn K, Caskey CT, Metzker M, Madellius C;
XX WPI; 1999-540560/45.
XX P-PSDB; AAY29953.
XX Human and mouse polynucleotides encoding CGICE polypeptides
XX Claim 2; Fig 2; 67pp; English.
XX
XX The present sequence represents the human CGICE cDNA sequence, which
XX when mutated is responsible for Best's macular dystrophy (BMD).
XX Polynucleotides encoding CGICE are useful for diagnosing whether a
XX patient carries a mutation in the CGICE gene. Normal and mutated
XX CGICE proteins are useful for identifying activators and/or inhibitors
XX of these proteins, in order to treat BMD. The CGICE gene offers a
XX simpler and cheaper method of diagnosing BMD without the need for the
XX presence of the patient. The gene may also be useful to discovering
XX the genetic cause of age-related macular dystrophy.
XX
XX Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 other;
XX
Alignment Scores:
Pred. No.: 7.73e-132 Length: 2229
Score: 1925.00 Matches: 366
Percent Similarity: 84.17% Conservative: 1
Best Local Similarity: 83.94% Mismatches: 0
Query Match: 81.91% Indels: 69
DB: 20 Gaps: 1
US-09-622-964-5 (1-435) x AA221227 (1-2229)
OY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 105 ATGACCATCACTTACACAAAGCCAAAGTGGCTATAGCCCGCTTAGCCTCTTCCCGCTG 164
OY 21 LeuLeuGlyTTPATGgLySerIleTyrIleuLeuTyrGlyIleuLeuIlePheLeu 40
DB 165 CTGGCTGCTCGCGGGGAGCATCAAGCTGCTATATGGCGAGTCTTAATCTTCTCG 224
OY 41 LeuGlyTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGlnGlnGlnLeu 60
DB 225 CTCGCTACATCAATCCGCTTATTTATAGGTGGCCCTCAGGAGAAACAAACAGCTG 284
OY 61 MethGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
DB 285 ATGTTTGAGAACTGACTCTGTATGGACAGCTACATCCAGCTCATCCCATTTCTCTTC 344
OY 81 ValLeuGlyPheTyrValThrLeuValValThrArgTyrTPRAsnGlnTyrGlnAsnLeu 100
DB 345 GTGCTGGGCTTCTACGTCAGCGTGTGTGACCCGCTGTGGAGACACAGTACAGAACTGTG 404
OY 101 PTCITTPRAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGln 120
|||||

```

```

DB 405 CCGTGGCCGAGCCGCTCATGAGCTGTGTGGGCTTCTGTGAAGCAAGACAGCA 464
OY 121 GlyArgLeuLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 140
DB 465 GGCCTGGCTGGCGGCGAGCTCATCCGCTACGCCCAACCTGGGCAACGTCCTCTCTG 524
OY 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnIleValGlnAla 160
DB 525 CGCAGCGTCAGCAGCCGAGCTCTTACAGCGCTTCCCGAGCGCCAGACCTGTGTGCAAGCA 584
OY 161 GlyPheMetThrProAlaGlnIleLysGlnLeuGlyLysLeuSerLeuProHisAsnMet 180
DB 585 GGCCTTATGACTCCGGAGAAACACACAGCTGTGAGAACTGAGCTACACACACACATG 644
OY 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyLysArg 200
DB 645 TTCGGTGGCTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 704
OY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgTyrGln 220
DB 705 ATCCGGGACCCCTATCCCTGCTCCAGACCTGCTGAAAGAGATGAAACACCTTGGCTACTCAG 764
OY 221 CysGlyHisLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnVal 240
DB 765 TGTGGACACCTGTATGCTTACAGCTGATGATGATGATGATGATGATGATGATGATGATG 824
OY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB 825 ACTGTGGCGGTGTACAGCTTCTCTGACTTCTAGTTGGCGGCACTTCTCAACCCA 884
OY 261 AlaLysAlaTyrProGlyHisGlnLeuAspLeuValValProValPheThrPheLeuGln 280
DB 885 GCCAGAGCTTACCCGCGCATAGAGCTGAGCTGCTGTGGCCGCTTACACTTCTCTCAG 944
OY 281 PhePhePheTyrValGlyTyrLeuLysValGlyLeuSerArgAlaLeuLeuGlyTyrPar 300
DB 945 TTCCTCTCTATGTTGGCTGCTCAAGG----- 972
OY 300 GHisGlyGlnArgGlyHisGlyGlnLeuLeuGlnThrArgMetGlnCysGlnGln 320
DB 972 ----- 972
OY 320 GlyValSerArgValGlnSerSerGlnAlaTyrTPRArgTyrProValIleProAla 340
DB 972 ----- 972
OY 340 PArgGlnAlaGlnGlyGlnSerLeuGlnProGlyArgArgArgLeuTyrTPRArg 360
DB 973 -----TGGCAGAG 980
OY 360 PSerSerSerThrProLeuGlnArgMetMetIleLeuArgProThrGlyLeuSerTh 380
DB 981 CAGCTCATCAACCCCTTGTGAGAGATGATGATGATGATGATGATGATGATGATGATG 1040
OY 380 rGlyIleCysArgCysProCysTyrPheTyrPheTyrGlySerThrArgTyrGlyLeu 400
DB 1041 AGCAATTTGGACAGTCTCCCTGTGGCTGTGATGATGATGATGATGATGATGATGATG 1100
OY 400 PSerArgThrCysThrGlyIleSerProSerHisSerProThrThrGlnLeuLeuPro 420
DB 1101 GAGCGGAGCATCTACTGGAATTAAGCCGAGACAGCCCTTACACAGCTGCTTCCGCC 1160
OY 420 cSerSerValIleProProLeuTyrPalaProProSerThrSerAla 435
DB 1161 CAGTTCGCTCGAGCCTCTTATGGGCTCCACCTTCAACATCAACGCC 1206

```

RESULT 3
AA221229 standard; cDNA; 1916 BP.

AA221229: 22-NOV-1999 (first entry)

ID ABA14558 standard; DNA; 18537 BP.
 AC ABA14558;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 6889.
 XX
 KW Human; nocotropic; neuroprotective; cytosolic; dermatological; virucide;
 KW immunosuppressive; antineoplastic; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antischizy; antianaemic; antiarthritic; cancer;
 KW antihumetic; hepatotropic; cerebroprotective; antinflammatory;
 KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 14-JUL-2000; 2000US-0217496.
 PR 26-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0225759.
 PR 22-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 23-AUG-2000; 2000US-0226868.
 PR 30-AUG-2000; 2000US-0227182.
 PR 01-SEP-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.

PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251150.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX Disclosure; SEQ ID NO 6889; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA14678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis; rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 other;
Alignment Scores:
Pred. No.: 6,39e-85 Length: 18537
Score: 1302.00 Matches: 383
Percent Similarity: 23.17% Conservative: 0
Best Local Similarity: 23.17% Mismatches: 1
Query Match: 55.40% Indels: 1269
DB: 22 Gaps: 6
US-09-622-964-5 (1-435) x ABA14558 (1-18537)
QY 51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluIuLysLeuThrLeuTyrCysasp 70
DB 4743 AGGCTGGCCCTCAGAGGAACAACAGCTGATGTTGAGAAACAGCTGTATTTGCGAC 4802
QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
DB 4803 AGCTACATCCAGCTCATCCCATTTCTCGTGGGTGAGTTCCCCCTTCTGGCTTT 4862
QY 82 ----- 82
DB 4863 CCGGGTCCCTGTGGCCCGCCAGGCTCCAGACAGGCCAGAGGAGATCAAGAGAGTGGC 4922
QY 82 ----- 82
DB 4923 GCAAGGGCTGGGAGGGGGGGGGAACGCCAGCGGAGGTGGCCCTCTCTGTAGGG 4982

QY 82 ----- 82
DB 4983 AAGGTGCGAGCTGCAGCCAGAGAAACTGAATTAGAGCTTAGTAAGACGTCTGCCGT 5042
QY 82 ----- 82
DB 5043 TAGCAATGAAGAACCCCATTTTCTGAGAGGAAGCGCTGACATCATGTCCTGAGCCCTTG 5102
QY 82 ----- 82
DB 5103 CCGGGAGGGAGGGGGTGTGCGGATTTCTGGACACAGAGGGGAGCCCGGGTGACA 5162
QY 82 ----- 82
DB 5163 GAACCTTGGGGCTCTCGGCTCCATGCGAGGCTCTGCTGCTCTGCCGAGCGC 5222
QY 82 ----- 82
DB 5223 CTTCCAGAGAGGCTGGGGGCTAGCCCGCTCCAGACAGAAAGCTGAGAGCGGACAT 5282
QY 82 ----- 82
DB 5283 CCGCGGGCGCTGGCCCTGGGCTCTGGCCGACGCTGGCCCTCGCCCTCGCCCGC 5342
QY 83 ----- GlyPheTyrValThrLeuValValThrArgTyrPheAsnIleTyrGlu 98
DB 5343 CCCTCTGCGCCAGGCTTGTAGCTGACGCTGCTGACCCGCTGGTGGACCACTAGACAG 5402
QY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluIleTyrAsp 118
DB 5403 AACCTGCGGTGCGCCGACCGCTCTATGAGCCGCTGGTGGGGCTGCTGGAAGGCAAGAC 5462
QY 119 GluGlnIleArgLeuLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
DB 5463 GAGCAAGCGCGGCTCTGCGGCGACGCTCATCCGCTACGCCAAGCTGGGCAAGCTGCTC 5522
QY 139 IleLeuAlaSerValSerThrAlaValIleTyrIleArgPheProSerAlaGlnHisLeuVal 158
DB 5523 ATCTGCGCAGCGTTCAGCACCGCAGTCTTCAAGCCCTTCCAGCGGCCAGCAGCAGCTG 5582
QY 159 Gln----- 159
DB 5583 CAGCAGGTGGGCGGACCGGAGCAACGGGAGGACCGGAGAGCCAGAGCGGCGGAGAT 5642
QY 159 ----- 159
DB 5643 GCGCGCGCAGGAATGGAAGATGGGTGAGCCAAAGTCCCGGACTCGGGGAGATTGGGT 5702
QY 159 ----- 159
DB 5703 GGAACCAAGAGTGGGGTGTGTCAAGATTGGGGGTCCAAATTGGGCGGAGAGATCGGG 5762
QY 159 ----- 159
DB 5763 TGTGTGAAGTGGGGCGAGCGAGAGACCCACCTCCGAGAGTAGAGCTGTAGAGCAGCG 5822
QY 159 ----- 159
DB 5823 CTAAGACCTTGAGGATATGAGAAAGAGGTGAGCGCTTGGGAAGTGTGAGTACT 5882
QY 159 ----- 159
DB 5883 AGGCTTACTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5942
QY 159 ----- 159
DB 5943 GGAATTTGTCTGTGACACCCCTCAGCGTGGGCTGACCTGTCCTGCTGCTGCTGCTGCTG 6002
QY 159 ----- 159
DB 6003 ACCCAGGCTAGGGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6062

Db	8223	ACAAACAACAACAACAACAAGCCCTAAAGTTACAGAGCCCTGCCCTTTAGAAAGCAG	8282
QY	357	-----	357
Db	8283	AGCGCACACACTCTCTCTTTATTCAGATGCGTGTGGCTGTCTTGTTCCTACCTCAAG	8342
QY	357	-----	357
Db	8343	TGGCTGTCCAGGATTCCCTCCACACAGCCAACTACGCGACAGATGTTCTGAATCA	8402
QY	357	-----	357
Db	8403	CACAGTTTCTCCACACTTTATCTTCTCTCTCTGTGTGCCACCACCTCTCTCTC	8462
QY	357	-----	357
Db	8463	CCTTCCTACCTCTTTATTTTTTGTGTAATGGGGGTGAAGTCTGTCTGTCCCTTC	8522
QY	357	-----	357
Db	8523	TGTCACTGTGACACACACACACACACACACACACACACACACACACTTCTAT	8582
QY	357	-----	357
Db	8583	TCCTTAAATTCCTCCCTGCCCTCCAGTTATCTTGTGTCTGCAGATCAAAACAATCA	8642
QY	357	-----	357
Db	8643	CACCTTTATGTTGAATTTCTCCAGGCTGCCCCAGTGGCTGCAGAAATGTCCCTGCAGC	8702
QY	357	-----	357
Db	8703	CCTAAGGCAGACGCGTGTACACTCTTCGGGCTTTGTAGGGCAVTTTAGAGTTGCTAT	8762
QY	357	-----	357
Db	8763	CCAGAACTGTGCCACACTAGACTGCCCCTTTGTTCACGCCACGTTCAATATATCTCTG	8822
QY	357	-----	357
Db	8823	TTGCATGAATGAATTAATTAATGCAACTCAGTAAGATACATGAGGTGATTAAGCA	8882
QY	357	-----	357
Db	8883	GTGACTACGCCGAGTGATACACTAGGACAGCTGTGGTTCAGGAGAGACTGCTC	8942
QY	357	-----	357
Db	8943	AGAAAGTTAGAGGGGCTGTGTCCAGAAAGTGTGGGTGCCCAAACTGTGGGGGCTGG	9002
QY	357	-----	357
Db	9003	AGCCTAAACTCTCCCTTTGAAGACAGTGTGAGGCAAGAGGGCTTCATGGGGTGTGA	9062
QY	357	-----	357
Db	9063	AATAGCAGACGTGAGGTTAAAGGGGGAAGCTGGCTTTGAGAGATTCCTCGAGGCTT	9122
QY	358	-----	358
Db	9123	TACAGAGCTCAGCTGTCCCAAAGTGGCGAGACCACTCATCAACCCCTTTGAGAGAGAT	9182
QY	369	lmetlleleuaurProthriclyleuSerthrglyleCys-----	383
Db	9183	GATGATATTTTGTAGACCACTGGAATGTGAGAGGAATTTTCAGGTATGGGGAAGGA	9242
QY	383	-----	383
Db	9243	GAGAAACCATCATGAGACCTTCCCAAAAGTGACCCAAAGAGAGAGACCACATGTCTG	9302
QY	383	-----	383
Db	9303	TAGGAGGCTCACAGTGAATGATCAACCCCTTCCCTCTCCCTCAGACAGCTCATTT	9362

Qy	383	-----	383
Db	9363	CACACAGAGATTCTACCTCAATCTTTGAGGCTCAGCAGGACACCATTCTCCCATTT	9422
Qy	383	-----	383
Db	9423	CACAGCAGGAAACTGAGGTCACAGAGAGAGAGATCTCCAAATCATCAGGCACA	9485
Qy	384	-----	385
Db	9483	TACAAGGTCCTGCCCTGGAGTATCTTCTGTGGACTTCTCTGTCCCTGTGACACAGT	9544
Qy	385	yspProCysTrpLeuTpmctArgCysThrArgThnCysIleuNlyTrpSerArgThrCysT	405
Db	9543	GTCGCCGTGTGGCTGTGTGATGAGATGCACCGACCTGCTCGGATGAGCGGACATGTA	9602
Qy	405	hcgIlyleSerProSerHIsSerProProThrGlnLeuLeuProProSerSerValGluP	425
Db	9603	CTGGAAATTAAGCCCCGAGCCACAGCCCCCTACACAGCTGCTCCGCCAGTTCCGTGCAGC	9663
Qy	425	roProLeuTrpAlaProProSerThrSer 434	
Db	9663	CTCCTTTATGGCTCCACCTTCAACATCA 9691	
RESULT 5			
ABAL4559			
ID	ABAL4559	standard; DNA; 16550 BP.	
XX	ABAL4559;		
XX			
AC			
XX			
DT	23-JUN-2002	(first entry)	
XX			
DE	Human nervous system related polynucleotide SEQ ID NO 6890.		
XX			
KW	Human; nocotropic; neuroprotective; cytostatic; dermatological; vituclide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary; antiparasitospinal; antistickling; antianemic; antilarthrific; cancer; antirheumatic; hepatocytic; cerebroprotective; antiinflammatory; antiatherogenic; antidiabetic; antilucer; anticonvulsant; antifungal; antiparasitic; cardiant; immune disorder; cardiovascular disorder; neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200159063-A2.		
XX			
PD	16-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01334.		
XX			
XX			
PR	31-JAN-2000; 2000US-0179065.		
PR	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		
PR	14-AUG-2000; 2000US-0224519.		
PR	14-AUG-2000; 2000US-0225213.		
PR	14-AUG-2000; 2000US-0225214.		

PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 30-AUG-2000; 2000US-0228294.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 26-SEP-2000; 2000US-0234998.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0242221.
 PR 08-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-02359678.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI, 2001-541565/60.
 Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 useful for preventing, diagnosing and/or treating nervous system
 cancers and metastases -
 Disclosure: SEQ ID NO 6890; 1701pp + Sequence Listing; English.
 The invention relates to novel genes (ABAI1004-ABA21534) and proteins
 (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
 medical conditions e.g. by protein or gene therapy. The genes are
 isolated from a range of human tissues disclosed in the specification.
 The nucleic acids, proteins, antibodies and (anti)agonists are useful
 in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 and ovarian cancer and other cancers of the adrenal gland, bone, bone
 marrow, breast, gastrointestinal tract, liver, lung or urogenital;
 (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 Note: The sequence data for this patent did not form part of the
 printed specification, but was obtained in electronic format directly
 from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 other;

Alignment Scores:

Pred. No.:	2,54e-84	Length:	16650
Score:	1293.00	Matches:	383
Percent Similarity:	23.20%	Conservative:	0
Best Local Similarity:	23.20%	Mismatches:	1
Query Match:	55.02%	Indels:	1267
DB:	22	Gaps:	6

US-09-622-964-5 (1-435) x ABA14559 (1-16650)

```
OY 51 ArgLeuAlaLeuThrGluGluGlnLeuMetPheGluLysLeuThrLeuTyrCysAsp 70
DB 4743 AGGCTGGCCCTCAGCGAAGAACACAGCTGATGTTGAGAAACTGACTGTATTGCGAC 4802
OY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
DB 4803 AGCTACATCCACCTCATCCCATTTCTTCGTGGTGAAGTACCTCCCTCTGCGCTTT 4862
OY 82 ----- 82
DB 4863 CCGGCTCCCTGTGGCCGCCAGGCTCCAGACAGGCCAGGAGATCACAGAGAGCTGCG 4922
OY 82 ----- 82
DB 4923 GCAGGGGCTGGGAGGGGGGGGAAACCCAGGGGCGAGTCCGCGCTCTGTAGGG 4982
OY 82 ----- 82
DB 4983 AAAGTGGGACTGACGACGAGAAACTGAACTAGAGTTAGTAAGACGTCTGCGGT 5042
OY 82 ----- 82
DB 5043 TAGCAATGAAAAACCCATTCTCTGAGGAGAGCGCTGACATCATGGTCCCTGAGCCCTG 5102
OY 82 ----- 82
DB 5103 CGCGGAGGGAGGGGCTCTGCGGATTTCTGGACACAGAGGGGACCCCGGGTGACA 5162
OY 82 ----- 82
DB 5163 GAACTTGGGGCTCTCCGCGCTCCATGCGAGGCTCTGCTGCTCTGCTCCGAGCGC 5222
OY 82 ----- 82
DB 5223 CTTCCAGAGAGGGGCTGGGGCTAGGCCCGCTCGCAGACAAAGCTGGAGGCCAGGCAT 5282
OY 82 ----- 82
DB 5283 CGCGGGGCGTGGGCCCTGGGCTCTGGCCGAGCGCTGGCCCTCGCCCTCGCCCGCGC 5342
OY 83 -----GlyPheTyrValThrLeuValValThrArgTyrTrpAsnGlnTyrGlu 98
DB 5343 CCCCTCCCTCCCAAGGCTTACGTGACGTGGCTGGACCCGCTGGTGAACCAAGTACAG 5402
OY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
DB 5403 AACCTGCGGTGGCCGACCGCTCATGTAGCCTGGGTGGGCTTCTGAAAGCCAGGAC 5462
OY 119 GlnGlnGlyArgLeuLeuArgArgThrLeuLeuArgTyrAlaAsnLeuGlyAsnValLeu 138
DB 5463 GAGCAGAGCGCGGTGGCTGGCGGACGCTCATCCGCTACGCCAACCTGGGACAGCTGCTC 5522
OY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal 158
DB 5523 ATCTGCGCAGCGTACAGACCGCAGCTACAAAGCGCTTCCCAAGCGCCACAGCTGTTG 5582
OY 159 Gln----- 159
DB 5583 CAACAGAGTGGCGGACCGGAGCAACGAGGAGGACCGGACAGAGCCAGGGCGCGAGAT 5642
OY 159 ----- 159
```

```
DB 5643 GGGCGGCGCAGCAATGGAAGATGGGTGAGACCCAAAGTCCCCCGGACTCGGGGATTGGGT 5702
OY 159 ----- 159
DB 5703 GGAGCCAGAGATGGGGTGTGTCAAGATTTTGGGGGTGCAATTGGGGGGGACAGAGTGGGG 5762
OY 159 ----- 159
DB 5763 TGTCTGAAGTGGGGGAGGCGCAGAGACCCACCTCCAGAGTAGAGTCTGAGCGAGGG 5822
OY 159 ----- 159
DB 5823 CTAAAGACCTTGAGGGATTAATGAAGAAGGTGACGGCTTGGAACTGTGAGGTACT 5882
OY 159 ----- 159
DB 5883 AGGCTACTACTCCCTCTGCCCCCTTGCCCCCTTGATCTCCGGTTTCCACTCTGAGGTATG 5942
OY 159 ----- 159
DB 5943 GGACATTTGTCTGTACACACCCCTCAGCCTGACCTGGTGGTCTGTAAATAGACAG 6002
OY 159 ----- 159
DB 6003 ACCCAGCCTAGCGCTGTGTGCTCTGCGCTGTAACTCCAGTGTCTTAGAGGCAAGGTGG 6062
OY 159 ----- 159
DB 6063 GAAATCGCTTGAGCCACAGCTGTGTAGACGCCCTGAGCAACATAGCAGACCCCATC 6122
OY 159 ----- 159
DB 6123 TCTACAAAAACATTAAAAATTAGCAGGCGAGTGGCGTGTGCTGTAGTCTGAGCGTGA 6182
OY 159 ----- 159
DB 6183 GTATCGGAGGCTGAGCGAGAGGATCACTTGAGCCAGCAATTCAGGCTGACGTGCGC 6242
OY 159 ----- 159
DB 6243 TAAATCGCACCGCTGACATCCAACTCGGTGACAGACAGACCCCTTCTCTGGAATA 6302
OY 159 ----- 159
DB 6303 AATTAATACCTGCCACATGCTCAGCCAGACAGCACCTAGTAGTGTCTGAGAAATT 6362
OY 159 ----- 159
DB 6363 TTTTGTGTGAAGAAAGAGGATGCAAAAGAGTGTGAGGTCTTATAGGTACAGAG 6422
OY 160 -----AlaGln 161
DB 6423 TGCGGCGCATCCCTTCTGACAGTTCCTCCACCCAGCGCTTCTTCACTCCACTCTGCA 6482
OY 161 yPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsnMetPhe 181
DB 6483 CTTTATGACTCCCGCAGAACCAAGCACTTGGAGAAATGAGCCCTACACACACATATTT 6542
OY 181 eTrpValProTyrValThrPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArgIle 201
DB 6543 CTGGGTGCCCTGGGGTGTGGTTGGCAACCTGTCAATGAAAGCGTGGGATGAGTCAAT 6602
OY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn----- 212
DB 6603 CCGGACCTTACTGTCTCCAGAGCTGTGTAAGTGAAGCCACTGTACAGACAGGCTGG 6662
OY 212 ----- 212
DB 6663 CCGCAGAGTGGGAAGGTTGTGTCTCACAGAAACAAGTTTCTACAAAGAAAGCCTT 6722
OY 212 ----- 212
```

```
D 6723 GGGCCCTGAGGCTCTCCGAGAGCCGAGGTGGGCTTGAGAACTTTTCCAAAGCA 6782
O 212 ----- 212
D 6783 TCCACAGCCCGAGGTGCTCCCTCTCAGAGGCCCTCCCTCTTCCAGTCTGTAGT 6842
O 212 ----- 212
D 6843 CCTGTTCCCTTTTGAATAGTAGAGAACTGAGACAAAGAGTTTGTAGCTTCCA 6902
O 212 ----- 212
D 6903 TGGCCACACAGCCAGGAATGACCATAGTACCAGGCCCTGTACTCGAAGAGGTGG 6962
O 212 ----- 212
D 6963 GGGCGAGCCCGAGGTGGGGGAGGTGTTGAGAACCCATCCCTCTTCCGCCCC 7022
O 213 --GluMetAsnThrLeuArgThrGlnCysGlyHisIleuTyraIlePheSer 232
D 7023 AGGAGATGAACACCTTGGCTACTCAGTGTGGACACTGTATGCTTGTGCTGATTAGTA 7082
O 232 IeProLeuValIleThr----- 237
D 7083 TCCCACTGGTGTATACACAGGTGAGGACTAGGCTGTGAGGCTGCTTTGGAACTG 7142
O 237 ----- 237
D 7143 AGGCTAGAAAGACCAAGAGACGCTGGGTGGAGGCTCAGCTAGAGCTAAGTGC 7202
O 237 ----- 237
D 7203 TCCCTGGAGTTGGGTCCACACTTGAAGTTGGGTGTGACTTTGAAGTGCAGTCT 7262
O 237 ----- 237
D 7263 AAGAGTCCAGGCTCCTGCTGGCCAGTCAGTAGAGCAATGTATATCCCATTTA 7322
O 237 ----- 237
D 7323 AAGAGAGTTGGCCGGGACAGAGTGGCTCATGCTGTATCCAGCACTTTGGAGCTGA 7382
O 237 ----- 237
D 7383 GGCAGGTGATCAGCTGAGGTCAAGAGTTCCAGACAGCTGGCCACATGTGAAACC 7442
O 237 ----- 237
D 7443 CATCTACTGAAATAACAGAAATTAGTGTGTGTGTGACAGGCTGTATCCAGCTAC 7502
O 237 ----- 237
D 7503 TTGGAGGCTGAGGACAGAGAAATGCTTGAACCCGGAGGTGAGGTTGCAGTGCTGA 7562
O 237 ----- 237
D 7563 GATCATGCACTGACATCCAGCCTGGGCGACAGCAAGACTGTCTCAAAACAACAAA 7622
O 237 ----- 237
D 7623 CAAACAACAACAACAAGGGTTAACAGAGCCCTAAGTACATAGTGTGCAAGTC 7682
O 237 ----- 237
D 7683 AGAACAAGGCTTGTCTCTCTCTCAGACTCCAGCCCTGGAGCATCTGTATTCAGG 7742
O 238 ----- 240
D 7743 GTTCCACCTAGCCCTTGTCTACCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7802
O 241 ThrValAlaValIlePheSerPhePheLeuThrCysLeuValAlaArgIlePheLeuAsnPro 260
D 7803 ACTGTGGCGGTGTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7862
O 261 AlaValAlaIleProGlnHisGlnLeuAspLeuValProValPheThrPheLeuGln 280
D 7863 GCCAAGGCTTAACCTGTGGCATGAGCTGAGCTGCTGTGTGCCCCGCTTCCAGTCTCTCAG 7922
O 281 PhePhePheThrValGlyTrpLeuValGlyLeuSerArgAlaLeuGlyTrpArg 300
D 7923 TTCTCTCTTAATGTGGCTGTGCTGAAGTGGGCTCTCCAGGGCCCTGTGGGCTGGAGG 7982
O 301 HisGlyGlnArgGlyHisGlyGlnLeuLeuGlnIleThrArgMetGlnCysGlnArg 320
D 7983 CATGCCAGAGGGGTATGCTATGCTCCAGCACTCTCCAGAGAGATGCACTGTCAAGAAAG 8042
O 321 LysValSerArgValAlaLeuSerSerGlnAlaIlePhePheArgThrProValIleProAlaThr 340
D 8043 AAGGTCTACGGGTAAAGACACAGCCAGGGGTGGCGCACACCTGTAAATCCAGCTACT 8102
O 341 ArgGluAlaGlyValIleGlyLeuSerLeuGlnProGlyArgArgArgLeuTrp----- 357
D 8103 CGGAGGCTGAGGACAGAGAAATCGTTGAACCGGGAGCGGAGTGTGTGATGAGT 8162
O 357 ----- 357
D 8163 ATCTGCCACTGCACTCCAGCTGGGCAAAAGATGAATCTATCTCAAAACAACAC 8222
O 357 ----- 357
D 8223 AACCAACAAACAAAGCCCTTAAGCTTCAGAAAGCCCTGCTTGAAGAGCAGACGGCAC 8282
O 357 ----- 357
D 8283 CACTCTCTCTTAATCAAGATGCTGTGGGCTGTCTGTCTCTCACTCACTCACTGCTGT 8342
O 357 ----- 357
D 8343 CCAGTAATCCCTCCACACAGCCAACTCCGAACAGATGTCTGAATACACAGTTT 8402
O 357 ----- 357
D 8403 TCTCTCACCTTATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8462
O 357 ----- 357
D 8463 CTTCTCTTATTTTGTGTAATGGGGGTGAAGTCTGTCTGTCTCTCTCTCTCTCTCT 8522
O 357 ----- 357
D 8523 TGACACACACACACACACACACACACACACACACACACACATCTCTCTCTCTCT 8582
O 357 ----- 357
D 8583 AAATTCCTCCCTGCCCCCACTTATCTTGTGTTCTGAGATCAAAACAATCACACTTT 8642
O 357 ----- 357
D 8643 TATGCTGAATTTCTCAAGGTGGCCCAAGTGGCTGCAAGATGTCCTGTGACCCCTAAG 8702
O 357 ----- 357
D 8703 GCAGAGCGGTGACCTCTTGGGGCTTGTGTAGGCGATTTAGAGTTGTGATTCAGGA 8762
O 357 ----- 357
D 8763 ATCTGCCACTGAGCTGCTCTTATGTTCAGCCAGCTTCAATATATCTCTGTTCAT 8822
O 357 ----- 357
D 8823 GAATGAATAAATATATGCAACTCCAGTAAGATACATGAGTGAAGTAAAGCAGTACT 8882
O 357 ----- 357
D 8883 CAGCCGATGATACACTCAGGAGACAGCTGTGGTGTTCAGGAAAGACTGTCTCAGAGA 8942
```



```
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239937.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0242221.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX
```

```
DR WPI: 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT using for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Disclosure; SEQ ID NO 6888; 1701bp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABAI1004-ABAI2134) and proteins
CC (ABAI678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 other;
Alignment Scores:
Pred. No.: 2,92e-84 Length: 18530
Score: 1293.00 Matches: 383
Percent Similarity: 23.20% Conservative: 0
Best Local Similarity: 23.20% Mismatches: 1
Query Match: 55.02% Indels: 1267
DB: Gaps: 6
US-09-622-964-5 (1-435) x ABAI4557 (1-18530)
QY 51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluLeuThrLeuTyCysAsp 70
DB 4743 AGGCTGGCCCTCAGGAGAAACACAGCTGATGTTGAGAACTGATGTTCCGAC 4802
QY 71 SerTyrIleGlnLeuIleProIleSerPheValIle----- 82
DB 4803 AGCTACATCCAGCTATCCCAATTCCTCGTGTGGAGTTCCTCCCTTGCGCTT 4862
QY 82 ----- 82
DB 4863 CCGGCTCCCTGCGCCGCCAGGCTCCAGACAGCGGAGGATCAGAGAGCTGCG 4922
QY 82 ----- 82
DB 4923 GCAGAGGGCTGGGAGGGGGCGGGGACCCAGCGGAGGTCCTCTGTAGGG 4982
QY 82 ----- 82
DB 4983 AAAGTGGGAGTGCAGCGCAGAGAACTGAGTGTAGGTAAGAGCTCTGCGCT 5042
QY 82 ----- 82
DB 5043 TAGCAATGAAACCCCAATTTCTGAGGAAAGCGTGCATGTCCTCGAGCCCTG 5102
QY 82 ----- 82
DB 5103 CGCGGAGGAGGAGGGGCTGTGGCGATTTCTGAGCAACAGAGGGGACCCCGGTGACA 5162
QY 82 ----- 82
DB 5163 GAACCTTGGGGCTTCGCGCTCCATCGAGAGGCTCTGCTGCTCGCGAGGCGC 5222
QY 82 ----- 82
```

Db	5223	CTTCCAGAGAGGCTGGGGCTAGGCCCTCGACAGCAAGCTGGAGAGCCGAGCAT	5282
Qy	82	-----	82
Db	5283	CGCCGGGCGTGGGCCCTGGGCTCGGCCGACGCTGGCCCTCGCCCTCGCCCGCCG	5342
Qy	83	-----	83
Db	5343	CCCTCGCCGAGGCTTCTACCTGACGCTGGTCGTGACCCGCTGGTGGCAACGATGACG	5402
Qy	99	AsnLeuProTPrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAsp	118
Db	5403	AACCTGCGGCGCCGACCGCTCATGAGCCGTGTGCGGCTTCGTGCAAGCAAGGAC	5462
Qy	119	GLUGInGlyArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu	138
Db	5463	GAGCAAGGCCGCTCTGGGCGGACGCTCATCCCTGACGCCACCTGGGCAACGTGCTC	5522
Qy	139	IleLeuArgSerValSerThrAlaValTyrIlyAspPheProSerAlaGlnHisLeuVal	158
Db	5523	ATCCTGCGCAGCGTCAGACCGCAGCTCTACAAAGGCTTCCCAAGGCCACCTGCTG	5582
Qy	159	Gln-----	159
Db	5583	CAAGCAGGTGGGCGGACCGGAGCAACGCGGAGCACCGGCGAGCCAGGGGCGAGAT	5642
Qy	159	-----	159
Db	5643	GGGGCGGCGAGGAATGGAAGATGGGTGAGGCCAAAGTCCCGGACTGGGGGAATTGGGT	5702
Qy	159	-----	159
Db	5703	GGAGCAGAGAGTGGGGTGTGTCAGATTGGGGGTCCAAATTGGCGGAGACAGATCGGG	5762
Qy	159	-----	159
Db	5763	TGTCTGAAGTGGGGCGAGCGCAGAGCCACCTCCGAGAGTAGAGTCTGAGCAGAGG	5822
Qy	159	-----	159
Db	5823	CTAAGGACCTTGAAGGATTAATGGAAGAAGGTGACGGCTTGGAACTGGTGAAGTACT	5882
Qy	159	-----	159
Db	5883	AGGGTCTACTTCCTCTGCCCCCTCTTATCTCGGTTTCCACTGAGAGTATG	5942
Qy	159	-----	159
Db	5943	GGACATTGGTCTGACACCCCTCAGCCTGGCTGACCTGGTCTGTTAATAAGACAG	6002
Qy	159	-----	159
Db	6003	ACCCAGGCTAGGGTGGTGGCTCTCGCTGTATCCAGTGTCTTAGAGAGCAAGGTGG	6062
Qy	159	-----	159
Db	6063	GAGAGTGCCTTAGACCCAGCTGTTTGAAGCCGCTGAGCAACATAGGAGAGCCCATC	6122
Qy	159	-----	159
Db	6123	TCTCAAAAACATTAAAAATTAGCAGAGGATGTGGCGTGTGCTGTAGTCTGAGGCTGA	6182
Qy	159	-----	159
Db	6183	GATCGGAGGCTGAGGAGAGGATCACTTAGGCCGACGATTCCAGGCTGCAGTGGC	6242
Qy	159	-----	159
Db	6243	TAAAGTGCACCGCTGCACTCAACTCGGTGACAGAGCAGACCTTTCTGTGAATA	6302
Qy	159	-----	159
Db	6303	AATTAATACCTGCCACATGCTCAGCCGAGACAGACCTAGTAGGTGCTCAGAAATTT	6362

Qy	159	-----	159
Db	6363	TTTTGTGTTAAGAAAGAGATGGCAAGAGATGCTGAGTTCCTATAGTCAACAGG	6422
Qy	160	-----	160
Db	6423	TGCCGGCATCCCTTCTGAGGTTCCTCCACCCGCGCTTCTTCACTCCACTTCGAGG	6482
Qy	161	-----	161
Db	6483	CTTATAGCTCCGCGAGAACACAGAGAGTGGAGAACTGAGCCTTACACAACTGTT	6542
Qy	181	eTTPValProTTPValTTPPheAlaAsnLeuMetIlyAlaTTPLeuGlyGlyArgT	201
Db	6543	CTGGGTGCGCTGGGTGTGGTTCCTCAACCTGTCAAGAAAGCGGTGGCTGGAGTGAAT	6602
Qy	201	eArgAspProIleLeuLeuGlnSerLeuLeuSn-----	212
Db	6603	CCGGACCTTATCTGCTCTCAGAGCTGTGAACGTGAGCCCACTGTACAGACAGGCTG	6662
Qy	212	-----	212
Db	6663	CCGAGAGTGGGAAGGTTGTGTCCACAGAAACAAAGGTTTCTCAAAAGAAAGCTT	6722
Qy	212	-----	212
Db	6723	GGGCCCCTGAGGGTCTTCCGAGAGCCGAGGTGGGGTTGCAGAACTTTCCAAAGCAA	6782
Qy	212	-----	212
Db	6783	TCCACAGCCGAGTGTGTCCCTTTCAGAGGCCCTCCCTTCTTCCAGTCTGTGAGT	6842
Qy	212	-----	212
Db	6843	CCTGTTCCTTGTATGATGAGAGAGTGAAGACAAAGAGTTTACTGAGCTTCCCA	6902
Qy	212	-----	212
Db	6903	TGGCCACACAGCCAGGAATGACCATAGTACCAGGCCCTGGTACCTGGAGAAAGGTGG	6962
Qy	212	-----	212
Db	6963	GGGCGAGCCAGAGGTGGGGGAGTGTGTTCAGAACCCCATCCCTCTTGGCCCCC	7022
Qy	213	--GluMetAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTPIleser	232
Db	7023	AGGAGATAAACACCTTGGTACTCAGTGTGACACCTGTATGCTTACGACTGGATTAGTA	7082
Qy	232	leProLeuValTyrThr-----	237
Db	7083	TCCACTGTGTATACACAGGTAGAGACTAGGCTGTGAGGCTTGGGAAACTG	7142
Qy	237	-----	237
Db	7143	AGGCTAAGAGACCAAGAGACAGCTGGGGTGGAGAGGCTCCTAGAGGCTAAGTGC	7202
Qy	237	-----	237
Db	7203	TCCCTGGAGTGGGTCCACACTTGAAGTTGGGTCTGAGACTTTGAAGTGCAGAGTCT	7262
Qy	237	-----	237
Db	7263	AAGAGTCCAGGCTCTGCTGGCCAGTCAGTACAGGCAATGTATTAATCCCATATTA	7322
Qy	237	-----	237
Db	7323	AAGAGAGTGGCGGCGACAGTGTGCTCATGCTGTAAATCCAGCACTTTGGGAAGCTGA	7382
Qy	237	-----	237
Db	7383	GCGAGGTGATCACTGAGGTTCAGAGAGTTTCAGAGACAGCCTGGCCAACTGTTGAACCC	7442

[illegible]

Db	8523	TGACACACACACACACACACACACACACACACACACACATCCTATTCTCT	8582
QY	357	-----	357
Db	8583	AAATTCCTCCCTGCCCCCAGTTATCTTTTGTTTCTGACATCAAAACAAATCACACTTT	8642
QY	357	-----	357
Db	8643	TATGCTGAAATTCCTCAGAGTGCCCCAGTGCCCTGCAGATGTCCTGGACCCCTAAG	8702
QY	357	-----	357
Db	8703	GCAGACGCGTGCACCTCTTCGGGCTTTGTTAGGCAATTTAGAGTTGCTATCCAGA	8762
QY	357	-----	357
Db	8763	ATTCGCCACCTAGACTGCCCTTATGTTACGCCACCTTCAGTATATCTCTGTTCAT	8822
QY	357	-----	357
Db	8823	GAATCAATAAATTTATGCACTCCAGGTAAAGATACATGAGGTGAGATAAAGCAGTACT	8882
QY	357	-----	357
Db	8883	CAGCGAGTGATACACTCAGGACAGCGTGTGGGTGTTCAAGGAAAGACTGCTCAGAGA	8942
QY	357	-----	357
Db	8943	GTTAGAGGGCTGTGTCCAGAAAGTGTGTGGTGCCCAAGATGTGGGGCGTGGAGCCCT	9002
QY	357	-----	357
Db	9003	AAACTTCCTTTGAAGACACTGGTCAGGCAAGAAAGGCTTCATGGGTGTGAATAAGC	9062
QY	357	-----	357
Db	9063	AGCAGCTGAGGTTTAAAGGGGAAGCTGGCTTTGAGAGCTCTGCCTAGAGGTTTACGA	9122
QY	358	-----TTPGLnSerSerSerThrProLeuGluArgMetLeu	371
Db	9123	GCCACACCTGCCCCAAGGTGGCAGACAGCTCATCAACCCCTTGGAGAGATGATGAT	9182
QY	371	titleArgProThrIleuSerThrGlyIleCys-	383
Db	9183	GATTTGAGCAACCTGATTTGTGCACAGAAATTTGCAAGGTATGGGAGAGAGAGAA	9242
QY	383	-----	383
Db	9243	CCATACCATGACCTTCCCAAGTGAACCAAGAGAGACCCACATGTTCTGTAGGA	9302
QY	383	-----	383
Db	9303	GGCCTCAGATGATATCAACCTTCCCTCCTCCTCCTCGACGATCATTCACCTCA	9362
QY	383	-----	383
Db	9363	CAGATTTCACCTCAATCTTTGAGGCTGAGGACAGCACCCATCTCCCATTTACACAG	9422
QY	383	-----	383
Db	9423	CAGGAAACTGAGGTCAGAGAGAGAGAGATTCCTCAAGTCATCAGGACATACAAAG	9482
QY	384	-----ArgCysProC	387
Db	9483	GMCTGCGCGGATATCTTTCTGTGGGACTTCTTGTCCCTGTCAGACAGGTGTCCTC	9542
QY	387	ySTPLeuTrpMetArgCysThrArgThrCysLeuGlyTrpSerArgThrCysThrGlyI	407
Db	9543	GTGGCTCTGATGATGATGACACAGACACTGCTCGGATGAGACCGGACATGTACTGGAA	9602
QY	407	leSerProSerHisSerProProThrGlnLeuLeuProProSerSerValGluProProL	427

```

Db      9603 TAAAGCCGAGCCACAGCCCCCTACACAGCTCTCTCCGCCAGTTCCGTGAGCCCTCTT 9662
Oy      427 eUTrPALaProProSerThrSer 434
        |||||
Db      9663 TATGGGCTCCACCTTCACATCA 9685

RESULT 7
AA221226
ID      AA221226 standard; DNA; 16125 BP.
XX
AC      AA221226;
XX
DT      22-NOV-1999 (first entry)
XX
DE      Human CG1CE genomic DNA sequence.
XX
KM      CG1CE; Best's macular dystrophy; mutation; diagnosis; detection;
XX      BMD; age-related macular dystrophy; ss.
XX      Homo sapiens.
XX
FH      Key
FH      Intron
FT      Location/Qualifiers
FT      1..1513
FT      /tag= a
FT      /number= 1
FT      1514..1570
FT      /tag= b
FT      /number= 1
FT      1571..3278
FT      /tag= c
FT      /number= 2
FT      3279..3466
FT      /tag= d
FT      /number= 2
FT      3315..15787
FT      /tag= e
FT      /product= "CG1CE protein"
FT      /note= "contains introns"
FT      3467..6616
FT      /tag= f
FT      /number= 3
FT      6617..6711
FT      /tag= g
FT      /number= 3
FT      6712..7227
FT      /tag= h
FT      /number= 4
FT      7228..7461
FT      /tag= i
FT      /number= 4
FT      7462..8353
FT      /tag= j
FT      /number= 5
FT      8354..8508
FT      /tag= k
FT      /number= 5
FT      8509..8896
FT      /tag= l
FT      /number= 6
FT      8897..8974
FT      /tag= m
FT      /number= 6
FT      8975..9676
FT      /tag= n
FT      /number= 7
FT      9677..10032
FT      /tag= o
FT      /number= 7
FT      10033..11212
FT      /tag= p
FT      /number= 8
FT      11213..11293
FT      /tag= q

```

```

FT      /number= 8
FT      Intron
FT      11294..11556
FT      /tag= r
FT      /number= 9
FT      11557..11708
FT      /tag= s
FT      /number= 9
FT      11709..13918
FT      /tag= t
FT      /number= 10
FT      13919..14557
FT      /tag= u
FT      /number= 10
FT      14558..15768
FT      /tag= v
FT      /number= 11
FT      15769..16125
FT      /tag= w
FT      /number= 11

W09943695-A1.
PD      02-SEP-1999.
XX
PF      22-FEB-1999; 99WO-US03790.
XX
PR      25-FEB-1998; 98US-0075941.
PR      18-DEC-1998; 98US-0112926.
XX
PA      (MERI) MERCK & CO INC.
PA      (UYUP-) UNIV UPSALA.
XX
PI      Petrukhin K, Caskey CT, Metzker M, Madellius C;
XX      WPI; 1999-540560/45.
DR      P-PSDB; AAY29953.
XX
DR      Human and mouse polynucleotides encoding CG1CE polypeptides
XX      P-PSDB; AAY29953.
XX
PS      Claim 2; Fig 1; 67pp; English.
XX
CC      The present sequence represents the human CG1CE gene, which when
CC      mutated is responsible for Best's macular dystrophy (BMD).
CC      Polynucleotides encoding CG1CE are useful for diagnosing whether a
CC      patient carries a mutation in the CG1CE gene. Normal and mutated
CC      CG1CE proteins are useful for identifying activators and/or inhibitors
CC      of these proteins, in order to treat BMD. The CG1CE gene offers a
CC      simpler and cheaper method of diagnosing BMD without the need for the
CC      presence of the patient. The gene may also be useful to discovering
CC      the genetic cause of age-related macular dystrophy.
XX
SO      Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 21 other:

Alignment Scores:
Pred. NO.: 1.57e-83 Length: 16125
Score: 1282.00 Matches: 383
Percent Similarity: 22.53% Conservative: 0
Best Local Similarity: 22.53% Mismatches: 1
Query Match: 54.55% Indels: 1317
DB: Gaps: 20

US-09-622-964-5 (1-435) x AA221226 (1-16125)
Oy      51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetHegLulysLeuThrLeuTyrCysAsp 70
        |||||
Db      6615 AGGCTGGCCCTCAGGAGAACACAGCTGATGTTGAGAAACGACTCTGTATTGCGAC 6674
Oy      71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
        |||||
Db      6675 AGCTACATCCAGCAGCATCCCATTTCTCTGCTGCTGAGTAGTCCCCCTTCGCGCTT 6734
Oy      82 ----- 82

```

```
Db 6735 CCGGGTCCCTGTGGCCGCCAGAGCTCCAGACAGGCCAGGAGATCACAGAGACTGCG 6794
QY 82 ----- 82
Db 6795 GCAGGGGCTGGGAGAGGGGCGGGGAAAGCCAGCGGCGAGTGGGGCTCTCTGTAGAG 6854
QY 82 ----- 82
Db 6855 AAAGGTGGGACTGCAGCCAGAGAAACTGAAAGTTAGACGTAGAGACGTCTGCCGT 6914
QY 82 ----- 82
Db 6915 TAGCAATGAAACCCATTCTGTAGGGAAGCGCTGACATCATGTCCCTGAGCCCTG 6974
QY 82 ----- 82
Db 6975 CGCGGAGAGGAGGGGGCTGTGGCGATTCTGGAGCAGACAGGGGAGCCCCGGGTGACA 7034
QY 82 ----- 82
Db 7035 GAACCTTGGGGCTCTGGCGCTCCATGCCAGGCTTGCTGCTCTGCCAGAGCG 7094
QY 82 ----- 82
Db 7095 CTTCAGAGAGGCTGGGGCTAGGCCGCTCGCAGCAGAAAGCTGAGAGCCGAGCAT 7154
QY 82 ----- 82
Db 7155 CGCGGGGCGCTGGGCCCTGGGCTGTGGCCGAGCGTGGCCCTCGCCCTCGCCCGCG 7214
QY 83 -----
          GlyPheTyrValThrLeuValValThrArgTyrPheProPheLeuValLeu
Db 7215 CCCTCTGCCCCAGAGCTTCTACGTAGCAGCTGTGTGTCGCCGCTGGTGAACCGTAGAG 7274
QY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAsp 118
          |||||
Db 7275 AACCTCCGTCGGCCACCGCTCATAGAGCTGTGGGCTTCGTCGAAAGCAAGAGAC 7334
QY 119 GluGlnGlyArgLeuLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
          |||||
Db 7335 GAGCAAGGCGGCTGTGTGGCGCCAGCGCTCATCCGCTAGCCCAACCTGGGCAACGTGCTC 7394
QY 139 IleuArgSerValSerThrAlaValTyrLysArgPheProSerIleGlnIleVal 158
          |||||
Db 7395 ATCTGTGGCAGCGTCTAGCACCGAGCTCTCAAGCGCTTCCGCCAGCCAGCACGTGGTG 7454
QY 159 Gln----- 159
Db 7455 CAAGCAGGTGGGCGGAGCCGGAGCAACGGGAGGCACCGGCGAGCCAGGGCCGAGAT 7514
QY 159 ----- 159
Db 7515 GGGCCCGGCAGAGAAAGAGATGGTGCAGCCAAAGTCCCGGACTCGGGGAGTGGGT 7574
QY 159 ----- 159
Db 7575 GGAGCAGAGAGTGGGGTGTGTGTCAGAGATTGGGGGTCCAAATTGGCGGGAGACGAGTGGG 7634
QY 159 ----- 159
Db 7635 TGTCTGAAGGTGGGCGAGGCCAGAGCCCAACCTCCAGAGTAGAGTCTGAGGCAAGG 7694
QY 159 ----- 159
Db 7695 CTAGGACCTTGTAGGAGTAATGAAAGAGGTGACGGCTTGGGAAGTGGTGAAGTACT 7754
QY 159 ----- 159
Db 7755 AGGGTCTACTTCCCTGTGCCCTTGTGATCTCCGTTTCCACTGTGAGATG 7814
QY 159 ----- 159
Db 7815 GGACATTGTCTGTACACCCCTCAAGCTGTGACTGATGTCCTGTGTTAATAAGACAG 7874
          |||||
QY 159 ----- 159
Db 7875 ACCAGGCTAGGCGTGTGGCTCTGCGCTGTATCCAGTGTATTAGAGGCAAGGTGG 7934
QY 159 ----- 159
Db 7935 GAAGATCCCTGTAGGCCAGCTGTGTGAGACGCCCTGAGCAACATAGCGAGACCCCATC 7994
QY 159 ----- 159
Db 7995 TCTACAAAACATTTAAATATTAGAGGGCATGTGGCGTGTCTGTAGTCTGAGGCTGA 8054
QY 159 ----- 159
Db 8055 GTATCGGAGGCTGAGCGAGAGGATCACTTAGGCCAGATTCCAGGCTGACGTGGCG 8114
QY 159 ----- 159
Db 8115 TAAGATGCACCGCTGCACCTCCAACTCGGTGACAGAGCCAGACCTTCTCTGGAAATA 8174
QY 159 ----- 159
Db 8175 AATAATACCTGCCACATGCTCAGCCAGACAGCACTAAGTGTGTGCAAAATT 8234
QY 159 ----- 159
Db 8235 TTTTGTGTTGAAAGAAAGAGATGCCAAAGAGTGTGAGTTCCTATAGTCAGCAG 8294
QY 160 -----
          AlaG 161
Db 8295 TGCCGGCATCCCTTCTGACAGTTCTCCACACCGCCTTCTTCACTCCACTGTGCA 8354
QY 161 yPheMetThrProAlaGlnHisLysGlnLeuGlyLysLeuSerLeuProHisAsnMetPh 181
          |||||
Db 8355 CTTTATGACACCTCCGAGAACACACAAAGAGTTGGAAACCTGACCTACACACAACTGTT 8414
QY 181 eTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyrLeuGlyArgGln 201
          |||||
Db 8415 CTGGGTGCCCTGGGTGTGTTTCCCAACCTGTAAATGAAGCGCTGTGGAGTGTGAAT 8474
QY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn----- 212
          |||||
Db 8475 CCGGAGCCTATCTCTCTCCAGAGCCTGCTGAACGTGAGCCACTGTACAGACAGGCTG 8534
QY 212 ----- 212
Db 8535 CCGCAGAGTGGGAAGGTTGTGGTCCACAGAAACAAGTTTCTCAAAAGAGAGCTT 8594
QY 212 ----- 212
Db 8595 GGGCCCTGAGGCTCTTCCAGAGCCGAGGTGGGTTGACAGAACTTTTCCAAAGCA 8654
QY 212 ----- 212
Db 8655 TCCAGAGCCGAGTGTGTCTTATCAGAGGCCCTCCCTCTTCTCCAGTCTGTAGGT 8714
QY 212 ----- 212
Db 8715 CCTGTCCCTTTGTATGATGAGAGAGCTGAGACACAAGAGTTTAGTAGCTTCCCA 8774
QY 212 ----- 212
Db 8775 TGGCCACAGACCCAGGAATGACCATAGTACAGGCCCTGTACTGTGAGAGAGGTGG 8834
QY 212 ----- 212
Db 8835 GGGCGAGCCAGGCTGGGGGAGGTGTGTTCAGAACCCATCCCTCTTCTGCCCCC 8894
QY 213 --GluMetAsnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTyrIleSer 232
          |||||
Db 8895 AGGAGATGAACACCTTGTGATCTGATGAGTGTGACACCTGTATGCTTACGACTGATTTGTA 8954
```

```
QY 232 leProLeuValTyrThr----- 237
Db 8955 TCCCACTGCTGTATACACAGGTGAGACTAGGCTGGTGGAGCTGCCCTTTGGGAAACTG 9014
QY 237 ----- 237
Db 9015 AGGCTAGAGGACCAGAGAGCAGCTGGGGTGGGAAGGCTCACCTAGAGGCTTAAGTGGC 9074
QY 237 ----- 237
Db 9075 TCCCTGGGAGTTGGGTCCACACTTTGAGTTGGTCTGGACTTTGAAAGTCCAAAGTTCT 9134
QY 237 ----- 237
Db 9135 AAGAGTCCAGGCTCTGCTGCGCCAGCTCCAGTAGAGGCAATGTGATTATCCCATTTA 9194
QY 237 ----- 237
Db 9195 AAGAGAGTTGGCGGCGCACAGTGGCTCATGCTGTAAATCCAGCACTTTGGGAAGCTGA 9254
QY 237 ----- 237
Db 9255 GCGAGGTGATCACTGAGTGCAGAGTTCCAGACCACTGGCCAAATGTAACCC 9314
QY 237 ----- 237
Db 9315 CATCTCTAGTAAATATACAGAAATTAGCTGTGTGTGTGACAGCCCTGTATCCAGCTAC 9374
QY 237 ----- 237
Db 9375 TTGGGAGCTGAGGACGAGAAATCGCTTGAAACCCGGGAGGTGAGGTGCAGTAGAGCTGA 9434
QY 237 ----- 237
Db 9435 GATCATGCGACCTGACTCCAGCCTGGGCGACAGCAAGCACTGTCTCAAAACAACAA 9494
QY 237 ----- 237
Db 9495 CAACAACAACAACAACAACAACAAGGGTTAACAGAGCCCTAAGTACATAAAGTG 9554
QY 237 ----- 237
Db 9555 TCGAAGTCAGAACAAAGCCCTTGCTCTCTCTCTCAGACTCCAGCCCTGAGACATCTG 9614
QY 238 ----- 238
Db 9615 ATTTAGGGTTCCCACTAGCCCTTTGCTAACACATCTCTCTCTCTCTCTCTCTCTCT 9674
QY 238 ----- 238
Db 9675 AGGTGTGACTGTGGGGGTGTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9734
QY 238 ----- 238
Db 9735 TGAACCCAGCCAAAGGCTTACCTGGCCATGAGACTGACTGTGTGGCCGCTTCTCTCT 9794
QY 238 ----- 238
Db 9795 TCTCTCAGTTCTTCTCTATGTGTGGCTGCTGAGAGTGGGCTCTCCAGGGCCCTGCTGG 9854
QY 238 ----- 238
Db 9855 GCTGAGGCGATGGCAGAGGGGTCTATGGCCAGCACTGCTTGAAGCGAGATGCGTCTC 9914
QY 318 lngluArgLysValSerArgValGluSerSerGlnAlaTPrpArgThrProValIleP 338
Db 9915 AAGAAAGAGAGGCTCTACGCGGTAGAAACAGCCAGCGCTGGGGCGACACCTGTAATCC 9974
QY 338 roaLeThrArgGluIuaGluIuaGluIuaGluIuaGluIuaGluIuaGluIuaGluIua 357
Db 9975 CAGCTACTCGGAGGCTAGAGGAGGAATGCTTGAACCCGGGAGGCTGTGTGTGTGTGTGT 10034
QY 357 ----- 357
Db 10035 GAGTGAAGATCGTGCCACTGCACTCCAGCCTGGGCAAAAGAAATGAATCTATCTCAAAA 10094
QY 357 ----- 357
Db 10095 ACAACAACAACAACAACAACAAGCCCTAAGTTTCAGAGGCCCTTGAAGACAGAGA 10154
QY 357 ----- 357
Db 10155 GCGAACACTCTCTATTAAGATGCTGTGGGTCTTTTCTACTCAGTAGCTGTCCAGTA 10214
QY 357 ----- 357
Db 10215 TTCTCCACAGCATATGAGAGATTTCTAATACAAATTTCTCACTCTAATTCCTCC 10274
QY 357 ----- 357
Db 10275 TTGTGCGACCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 10334
QY 357 ----- 357
Db 10335 CTGCATTGAGACAAAATACAGAGAGAGAGAAAGATCTATCTAATCCGCCCATTTAG 10394
QY 357 ----- 357
Db 10395 TTGAAAAAAACTTATTAAATCAGGCAAGTAAATCCGCAAGATTGNNNNNNNNNA 10454
QY 357 ----- 357
Db 10455 GATGTTCTGAATCAGAGAGTTTCTCTGAGCTTTTATCTTCTCTCTCTCTCTCTCT 10514
QY 357 ----- 357
Db 10515 ACCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10574
QY 357 ----- 357
Db 10575 GTCTGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10634
QY 357 ----- 357
Db 10635 ACACACATTCCTATTCCTAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10694
QY 357 ----- 357
Db 10695 ATGTCCTGGACCCCTAAGGAGAGCGGTGTCACTCTTGGGGCTTTGTTAGGCATT 10814
QY 357 ----- 357
Db 10815 TTAGAGTTGCTATTCAGGAATCTGCCACCTAGACTGCTTATGTTCAAGCCAGCTTC 10874
QY 357 ----- 357
Db 10875 AGTATATATCTCTGTGATGATGAATGAATGAATGAATGAATGAATGAATGAATGAG 10934
QY 357 ----- 357
Db 10935 GTGAGATPAAAGCAGTACGCTCAGCCAGGTATACACTCAGAGAGAGCTGGGTGTTCA 10994
QY 357 ----- 357
Db 10995 GGAAGACTGGCTCAGAAAGATTAGAGGGCTGTGTCAGAAAGTGTGGTGCCTACAA 11054
QY 357 ----- 357
Db 11055 GTGTGGGGCTGGAGCCCTTAATCTGCTTGAAGACAGTGTCAAGGAGGAAGGCT 11114
QY 357 ----- 357
```

Db 11115 TCATGGGTGTGAAATAGCAGCAGTGTAAAGGGGAAGCTGGCTTGAGGAGT 11174
Qy 358 -----TrrGlnSerSerSerThr 364
Db 11175 TCGCCGTGAGGGTTTACAGAGCCTCACCTGTCCCAAGTGGCAGAGACACTATCAC 11234
Qy 365 ProLeuGluArgMetMetIleLeuArgProThrGlyLeuSerThrGlyIleCysArg 384
Db 11235 CCTTGGAGAGATGATGATGATTTTGACCAACTGATTTGTGACAGGAATTTGCAG- 11293
Qy 384 -----
Db 11294 GATGGGAGAGGAGAGAAACCATATCATGACCTTCCCAAGTGGACCCAAAGAGAG 11353
Qy 384 -----
Db 11354 CTCCTCCCTCTGACAGCCAGTATTCATCTACAGATTTCTACCTCATCTTTGAGCGTG 11413
Qy 384 -----
Db 11414 CAGCAGACACCCATCTCCCATTTACAGCAGGAAAGTGGTCCAGAGAGAGGAG 11473
Qy 384 -----
Db 11474 AGATTCTCCAGTATCATGAGCAGCATACAGAGTCCGCTGGATGATCTTCTGGGA 11533
Qy 385 -----CysProCysTrpLeuTyrMetArgCysThrArgThr 396
Db 11534 CTTCCTTCCTCCCTGGTACACAGGAGTCCCTGTGGCTGTGGATGATGATGACACAGAG 11593
Qy 397 CysLeuGlyTrpSerArgThrCysThrGlyIleSerProSerHisSerProThrGln 416
Db 11594 TGCCCTCGATGAGACCCGAGCATGTAGTGAATTAAGCCGACAGCCCTTACACAG 11653
Qy 417 LeuLeuProProSerSerValGluProProLeuTrpAlaProProSerThrSer 434
Db 11654 CTGCTTCGCGCCCACTCCGTGACCTCTTATGAGGCTCCACTTCACATCA 11707
RESULT 8
ID ABA14556
XX ABA14556 standard; DNA; 7108 BP.
AC XX
XX ABA14556;
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6887.
KW Human; nootropic; neuroprotective; cytoskeletal; dermatological; viroclade;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiParkinsonian; antiskilling; antianemic; antiarthritic; cancer;
KW antileukemic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiact; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
PD 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01334.
PF
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225275.
PR 14-AUG-2000; 2000US-0225279.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SW;
 XX WPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure: SEQ ID NO 6887; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AB11004-ABA21534) and proteins
 CC (AB114678-AB18001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 7108 BP; 1632 A; 1882 C; 1989 G; 1605 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,8e-79 Length: 7108
 Score: 1216.00 Matches: 347
 Percent Similarity: 24.05% Conservative: 0
 Best Local Similarity: 24.05% Mismatches: 5
 Query Match: 51.74% Indels: 1094
 DB: 22 Gaps: 5
 US-09-622-964-5 (1-435) x AB114556 (1-7108)
 QY 83 GlyPheTYrValThrLeuValValThrArgTTPTPAsngIntyGluAsnLeuProTP 102
 DB 76 GGCTTCTAGCTAGCGGTGCTGAGCCCTGCTGAGACCACTAGAGAACTGCCGTG 135
 QY 103 ProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspLungInglArg 122
 DB 136 CCCGACCGCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 195
 QY 123 LeuLeuArgArgThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 142
 DB 196 CTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 255
 QY 143 ValSerThrAlaValTyrTyrAspPheProSerAlaGlnHisLeuValGlnAla 160
 DB 256 GTGACGACCGGCGGCTGACAGCGCTTCCCGACCGCGGCGGCGGCGGCGGCGGCGGCGG 314
 QY 160 160
 DB 315 GCGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 374
 QY 160 160
 DB 375 GAATGAGATGGGTGAGGCCAAAGTCCCGGAGCTGGGGGATTGGGTGGAGCCAGCA 434
 QY 160 160
 DB 435 GTGGGTGTAGTCAAGATTGGGGGTCCAAATTGGGCGGAGACGATCGGTGTGAAG 494
 QY 160 160
 DB 495 TGGGCGGAGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 554
 QY 160 160
 DB 555 TTGAGGATATGGAAGAAGGTGACGGCTTGGGAACGTGAGTACTAGAGGTACTACT 614
 QY 160 160
 DB 615 TCCTCTGCGCTTGGCCCTCTTGATCTCCGGTTTCCACTCTGAGGATATGGGACATTGGT 674
 QY 160 160
 DB 675 CTCTGACACCCCTCAGCTGCGCTGACCTGTGCTCTGTTATATAGACAGACCCAGCTA 734
 QY 160 160
 DB 735 GCGGTGTGCTGTCCCTGTATATCCAGTGTCTTATAGAGCAAAAGGTGGGAAGATCGCT 794


```

OY 160 ----- 160
Db 795 TGAGCCAGCTGTTTGAGACGCCCTGAGCAACATAGCGAGACCCCATCTCTACAAAA 854
OY 160 ----- 160
Db 855 CATTAAAAATTAGCAGGCGCATGTGGGTGTGCTGCACTGTGAGGCTGAGTATCGGAG 914
OY 160 ----- 160
Db 915 GCTAGGCGAGAGATCATTGAGCCAGCAGATTCCAGGCTGAGTGGCTAAGATCCGA 974
OY 160 ----- 160
Db 975 CCAGTGCATCCAACTCGGTGACAGAGCCAGACCTCTCTGGAATAATAAATATAC 1034
OY 160 ----- 160
Db 1035 CTGCCACATGCTCAGCCAGAACAGCAGCTAGTAGGTGCTCAGAAATTTTTTTGTT 1094
OY 160 ----- 160
Db 1095 GAAAGAAAGAGATGGCAAGAGAGTGTGAGTTCTATAGTCAGCAGCGCGCCAT 1154
OY 161 ----- 161
Db 1155 CCCTCTGAGGTTCTCCACCCAGCCCTCTTCACTCCACTCTGCAAGGCTTATAGACT 1214
OY 165 ProAlaGluHisIysGlnLeuGluIlyLeuSerLeuProHisAsnMetPheTrpValPro 184
Db 1215 CCGGCAACACAAAGAGAGTTGGAAAGCTGAGCCCTACACACAACTGTTCTGGGTGCC 1274
OY 185 TrpValTrpPheAlaAsnLeuSerMetIlySalaTrpLeuGlyIlyTrpIleArgAspPro 204
Db 1275 TGGGTGTGGTTGGCCAACTGTCAATGAAGGCTGGCTTGAGGTCGAATCCGGGACCT 1334
OY 205 IleLeuLeuGlnSerLeuLeuAsn ----- 212
Db 1335 ATCTCGCTCCAGAGCTGCTGAACGTGAGCCCACTGTACAGACAGGCGTCCGCAAGTG 1394
OY 212 ----- 212
Db 1395 GGAAGGCGTGTGCTCCACAGAAACAGGTTTCTACAAAGAGAAAGCTTGGCCCTGAG 1454
OY 212 ----- 212
Db 1455 GGTCTCCGAGAGCCTGAGGTGGGTGTCAGAAATTTTCCAAACAGCAATCCACAGCCG 1514
OY 212 ----- 212
Db 1515 AGGTGGTCCCTTCTCAGAGGCCCTCCTCTTCTCAGAGTCTGTAGAGTCCGTGCCCT 1574
OY 212 ----- 212
Db 1575 TTTGATAGATAGAGAGCTGAGACACAAAGAGTTTATGAGCTTCCATGGCCACAG 1634
OY 212 ----- 212
Db 212 ----- 212
OY 213 ----- 213
Db 1635 CCAAGAAATGACATAGGTACAGAGCCCTGTACTGTGAGAGAGGTGGGGCGAGCCCA 1694
OY 213 ----- 213
Db 1695 GGGTGGGGGAGGTGTGTTCCAGAACCCCATCCCTCTTCTGCCCCCGACAGAGATGAAAC 1754
OY 216 ThrLeuArgThrGlnCysGlyHisLeuTyraIleTrpAspTrpIleSerIleProLeuVal 235
Db 1755 ACCTGGTACTCAGTGTGACACCTGTATGCTTACAGCTGAGTGAATATCCACTGTGT 1814
OY 236 TyrThr ----- 237
Db 1815 TATACACAGGTGAGACTAGCTGTGTGAGGCTGCCCTTTTGGAAACTGAGGCTAGAGG 1874
OY 237 ----- 237

OY 1875 ACCAAGAGAGAGCTGGGGTGGGAAGGCGCTACCTAGAGGCTAAGTGCTCCCTGGAG 1934
OY 237 ----- 237
Db 1935 TTGGGTCCACACTTTGAAGTTGGGTCTGCACTTTGAAGTGCAGAGTTCAAGTCAG 1994
OY 237 ----- 237
Db 1995 CTCTGCTGCGCCAGTCCAGTAGAGCAATGTGATTAATCCCATATTAAAGAGTTG 2054
OY 237 ----- 237
Db 2055 GCCGGTGCAGTGGCTCATGCTGTAAATCCAGCACTTTGGAGAGTGAAGAGTGAT 2114
OY 237 ----- 237
Db 2115 CACCTGAGTCAAGAGTTGAGAACCAACCTGGCCACATGAGTGAACCCCATCTACTG 2174
OY 237 ----- 237
Db 2175 AAAATACAGATTAGCTGTGTGTGTGTCACGCTGTAAATCCAGTACTTTGGAGGCTG 2234
OY 237 ----- 237
Db 2235 AGGCAAGAGAAATGCTTGAACCCGGAGGTGAGGTTGAGTGAAGTATCATGCCAC 2294
OY 237 ----- 237
Db 2295 TGCACCTCAGCCTGGGCGACACAGAAAGTGTCTCAAAACAAACAAACAAACAA 2354
OY 237 ----- 237
Db 2355 ACAAAAGGTTTAAACAGAGCCCTTAAGTACATAGTGTCAAGTCAAGAACAGGCTTG 2414
OY 237 ----- 237
Db 2415 TCTCTGTCTAGACTCCAGGCCCTGGAGCATCTGATTTAGAGGTTCCACCTAGCCC 2474
OY 238 ----- 238
Db 2475 TTTGCTACCACTCTCTCTCTCTCTCCAGAGTGTGAGTGTGGCGGTGACAGCTTCTTC 2534
OY 249 LeuThrCysLeuValGlyArgGlnPheLeuAsnProAlaIlySalaTrpProGlyHisGlu 268
Db 2535 CTGACTGTCTAGTTGGCGCGAGTGTCTGAACCCAGCCAGGCGCTACCTGGCCATGAG 2594
OY 269 LeuAspLeuValIleProValPheThrPheLeuGlnPhePhePheTrpValGlyTrpLeu 288
Db 2595 CTGGACCTCGTTGTGGCCGCTTTCACGTTCTCTGCACTTCTTCTATGTGGCTGGTG 2654
OY 289 LysValGlyLeuSerArgAlaLeuLeuGlyTrpArgHisGlyGlnArgGlyHisGln 308
Db 2655 AAGGTGGCCCTCTCAGGGCCCTCTGCGCGGAGGAGCATGGCCAGAGGGGTATGGCCAG 2714
OY 309 GlnLeuLeuGluThrArgMetGlnCysGlnIlyArgIlySalaIlySerArgValGlySer 328
Db 2715 CAGTGGCTGTGAGAGAGAGAGTGTGAGGAAAGGAGGCTCTACGGGTGAAGAGCAGC 2774
OY 329 GlnAlaTrpTrpArgThrProValIleProAlaThrArgGlnIlyArgIlyGlySer 348
Db 2775 CAGCGTGTGGGCGCACACCTGTAAATCCAGCTACTGGGAGGCTGAGGCGAGGAATCG 2834
OY 349 LeuGlnProGlyArgArgGlyLeuTrp ----- 357
Db 2835 CTTGAAACCGGAGGCGGAGGTTGTGTGATTGAGATCGTGCACCTGCACCTCAGGCTG 2894
OY 357 ----- 357
Db 2895 GGCAGAAAGATGAACCTATCTCAAAACACAAACAAACAAACAAAGCCCTAAGGT 2954
OY 357 ----- 357

```

Db 2955 TCAGAAAGCCCTGCTTTAGAAAGCAGATCGGACACGCTCTCTATTCAAGATGCT 3014
 QY 357 ----- 357
 Db 3015 GTTGGGCTGCTTGTTCCTCACCCTCAGTGGCTTTCAGGTAATTCCTCCACACAGC 3074
 QY 357 ----- 357
 Db 3075 CAATACCCGACAGATGTTCTGAATACACAGTTTCTCTCCACCTCTTATCTTCTTCT 3134
 QY 357 ----- 357
 Db 3135 TCCTTGTGTCACCACT 3194
 QY 357 ----- 357
 Db 3195 GGGTGAAGTCTCTGTCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3254
 QY 357 ----- 357
 Db 3255 ACACACACATACACACACACACACACGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3314
 QY 357 ----- 357
 Db 3315 CCCAGTTATCTTGGTTCTGAGATCAAAACAATCACACTTTATGCTTGAATTCCTC 3374
 QY 357 ----- 357
 Db 3375 CGGGGTCGCCAGTGGCTGCAAGATGTCCTGACCCCTTAAGCAGACGCTGTACAC 3434
 QY 357 ----- 357
 Db 3435 TCTTGGGGCTTGTGAGGCAATTTAGAGGTTGCTATCCAGAAATCTGCCACCTAGAC 3494
 QY 357 ----- 357
 Db 3495 TGCCTTAGTTCAGCCAGCTTCACTATATCTCTGTGATGATGAATTAATTAAT 3554
 QY 357 ----- 357
 Db 3555 GCACTCCAGTGAAGATACATGAGTGAATTAAGCAGTACTCAGCCGATGATACAC 3614
 QY 357 ----- 357
 Db 3615 TCAGGACAGCTGTGGGTTCAGGGAAGAGCTGCTCAGAAAGATTAGAGGGCTGTGT 3674
 QY 357 ----- 357
 Db 3675 CCAGAGTGTGTGGTGCACCAAGTGTGGGGCTGAGCCCTTAACCTGCTTTGAGAG 3734
 QY 357 ----- 357
 Db 3735 CAGTGTACGACGAGAAAGGCGTCATGGGCTGTGAATAAGCAGCTGAGGTTAAAG 3794
 QY 357 ----- 357
 Db 3795 GGGGAAGCTGCTTGAAGAGTTCGCTGAGGGTTTACAGAGCCTCACCTGCCCAAG 3854
 QY 358 -TTPGInserSerSerSerThrProLeuGluArgMetMetIleLeuArgProThrI 377
 Db 3855 GTGGCAGACACTCATCAACCCCTTGGAGAGATGATGATTTTGAAGCAACTCG 3914
 QY 377 yLeuSerThrgIyIleCyArg----- 384
 Db 3915 ATTGTGCACAGAAATTTTCAG-CTATGGGAGAGAGAGAAACCATACATGAGACTTC 3973
 QY 384 ----- 384
 Db 3974 CCCAAAGTGACCAAAAGAGAGACCCACTGTTCGTAGGAGGCCCTCAGAGTGAATGA 4033
 QY 384 ----- 384
 Db 4034 TCACCCCTTCCCTCCTCCTCCTCCTCAGCAGCATTCATCAGCAGGATTCACCTCA 4093

QY 384 ----- 384
 Db 4094 TCTTTGAGCTGAGCAGCAGCACCACCATCTCCCATTTACAGCAGGAGAAATGAGTCC 4153
 QY 384 ----- 384
 Db 4154 AGAGAGGAGAGAGATTCTCCAAATCATCAGGCACATACATAGTCTTCTGGATGA 4213
 QY 385 ----- 385
 Db 4214 TCTTCTGTGGACTTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4273
 QY 392 gCysThrArgThrCysLeuGlyTrpSerArgThrCysThrgIyIleSerProSerHis 412
 Db 4274 ATGCACACAGCAGCTGCTCGATGAGAGCCGACATGATGTAATAGCCGACACAG 4333
 QY 412 rProProThrgIleuLeuProSerSerValGluProLeuTrpAlaProProse 432
 Db 4334 CCCCCCTACACAGCTGTT-CCGCCAGTTCGTCGAGCCCTCTTATGGCTCCACCTTC 4392
 QY 432 rThrSer 434
 Db 4393 AACATCA 4399
 Db 4393 AACATCA 4399
 RESULT 9
 ABL10793
 ID ABL10793 standard; cDNA; 2861 BP.
 AC ABL10793;
 XX
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; seq.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EM;
 DR MPI: 2001-656860/75.
 DR P-PSDB; ABB66690.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 PS
 PS Claim 1; SEQ ID NO 26861; 21bp + Sequence Listing; English.
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2861 BP; 761 A; 701 C; 723 G; 676 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.39e-49 Length: 2861
 Score: 801.00 Matches: 183
 Percent Similarity: 54.09% Conservative: 55
 Best Local Similarity: 41.59% Mismatches: 107
 Query Match: 34.09% Indels: 95
 DB: 23 Gaps: 9
 US-09-622-964-5 (1-435) x ABL10793 (1-2861)
 QY 1 MethrIleThrTyrThrSerGlnValAlaAsnAlaArg---LeuGlySerPheSerArg 19
 DB 415 ATGACAAATTCACACAGTGAAGTGAAGTGGCCACTTGGCGCTTGGCTGTTCACAA 474
 QY 20 LeuLeuLeuCysTrpArgGlySerIleTyrLysLeuLeuTyrGlyGluPheLeuIlePhe 39
 DB 475 TTGCTGCTCAGATGGCGAAGACATTTACAAACGTGGCTTGGCTAGATCTTCTGGCCTTC 534
 QY 40 LeuLeuCysTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGluGluGlnGln 59
 DB 535 TTGACCAATTTACTATGCGATCAACATGGTGTATCGCTTGGCCTCAACCCGCAAA 594
 QY 60 LeuMetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSer 79
 DB 595 GAACCTTTGAGGCCATTTCTCAGTCTGATGTATGATGACCCCTGGTGAATCAGTACCTCC 654
 QY 80 PheValLeuGlyPheTyrValThrLeuValAlaThrArgTyrPheAsnGlnTyrGluAsn 99
 DB 655 TTCGTGCTTGGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 714
 QY 100 LeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGlu 119
 DB 715 ATTCCTGGCCAGATCCCATCGCGTGTTCACAGCTGGAATGTCATGATGATGATGATG 774
 QY 120 GlnGlyArgLeuLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIle 139
 DB 775 CGAGACGACATGATGAGCGGCAACATATGCGATGTGTGCTTGGCTGACGATGATG 834
 QY 140 LeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGln 159
 DB 835 CTGGCGAATGTTCGCGAGGAGTGAAGAAGCGTTTCCCGCCCTAAATATCTGTGTGAA 894
 QY 160 AlaGlyPheMetThrProAlaGluHisLysGlnLeuGluLysLeu-----SerLeuPro 177
 DB 895 GCGGGTCTGCTAATGACAAATGAAGAACCATCATGAGACCATGAACAAGGCTTCC 954
 QY 178 His---AsnMetPheThrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTyr 196
 DB 955 AGACCTTCGAAGCACTGGCTGCCATCGTGTGGCTGCCCATATTAACCAAGCGCCAGA 1014
 QY 197 LeuGlyGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThr 216
 DB 1015 AAGGAAGGTGCGATTCGTGATGATTTGCTGTAAGACCATCATGAGCTAAATTAAG 1074
 QY 217 LeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyr 236
 DB 1075 TTTCGTGGTCAAGTGTGAGTCTCATCAGCTACATACATTAAGTACCTGTGTGTAC 1134
 QY 237 ThrGlnValValThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGln 256
 DB 1135 ACCCAAGTGTGACCCCTGGCGGTGTATTCGTACTCTCTTACCTGCTGCTGCAACAA 1194
 QY 257 -----PheLeuAsnProAlaLysAlaTyrPro 265
 DB 1195 TGGACCGATGCGAAGGTGTGGCGAATACCACTACTGTAAC----- 1236
 QY 266 GlyHisGlnLeuAspLeuValValProValPheThrPheLeuGlnPhePheTyrVal 285
 DB 1237 -----AAGGTGATCTATACTTTCCTGTATTACAAAGCTGAGGTCTCTTCTACATG 1290

QY 286 GlyTyrPheLysValGlyLeuSerArgAlaLeuLeuGlyTyrArgHisGlyAlaArgGly 305
 DB 1291 GGTGGCTCAAGTGT----- 1305
 QY 306 HisGlyGlnGlnLeuLeuGluThrArgMetGlnCysGlnGluArgLysValSerArgVal 325
 DB 1305 ----- 1305
 QY 326 GluSerSerGlnAlaTyrTrpArgThrProValIleProAlaThrArgGluAlaGluAla 345
 DB 1305 ----- 1305
 QY 346 GlyGluSerLeuGluProGlyArgArgLeuTyrTrpGlnSerSerSerThrPro 365
 DB 1306 GCGGATCTCGTG-----ATTAATCA 1326
 QY 366 -LeuGluArgMetMetMetIleLeuArgProThrGlyLeuSerThrGlyIleCysArgCys 385
 DB 1327 TTTGGCGAAGACATGATGATTTTGAAGTCACTGATGTGTGATCGCATCTTCAAGTG 1386
 QY 385 sProCysTrpLeuTyrMetArgCysThrArgThrCysLeuGlyTyrSerArgThrCysTh 405
 DB 1387 TCCTATGATGATGTCGACGAGATGACACATGACCATCGGAGCTTTAAAGATCACTAC 1446
 QY 405 rGly-----IleSerProSerHisSerProProThrGlnLeuLeuProProSerSer 422
 DB 1447 TGGACAGAGGTGTCCCAACG---AGCTGCCCTACACATACTCTGCCAAGCATTC 1501
 RESULT 10
 ABL12609
 ID ABL12609 standard; cDNA; 1608 BP.
 AC ABL12609;
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32309.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 PN WO200171042-A2.
 XX 27-SEP-2001.
 PD 23-MAR-2001; 2001WO-US09231.
 PF 23-MAR-2001; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 PR (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 PL P-PSDB; ABB68506.
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Claim 1; SEQ ID NO 32309; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC

xx Sequence 1608 BP; 431 A; 336 C; 419 G; 422 T; 0 other;

Alignment Scores:

Pred. No.:	5,276-41	Length:	1608
Score:	685.00	Matches:	126
Percent Similarity:	65.31%	Conservative:	66
Best Local Similarity:	42.86%	Mismatches:	98
Query Match:	29.15%	Indels:	4
DB:	23	Gaps:	2

US-09-622-964-5 (1-435) x ABL12609 (1-1608)

```

OY 1 MethrILethrYrThrSerGlnValAlaAsn--AlaArgLeuGlySerPheSerArg 19
DB 1 ATGACTGCTCTTACACCGCTGAGGTGCGACATGACCATTCGGGCTTCGGAAG 60
OY 20 LeuLeuGlySerPheGlySerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 39
DB 61 CTTTGATGAGATGGCGCAGATATCAAGATAAATGGTGATCTTCTGCAATTC 120
OY 40 LeuLeuGlySerPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 59
DB 121 CTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
OY 60 LeuLeuGlySerPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 79
DB 181 CCTGTTTCGAGCATGTGATGATGATGATGATGATGATGATGATGATGATGATG 240
OY 80 PheValLeuGlyPheLeuValLeuValLeuValLeuValLeuValLeuValLeu 99
DB 241 TTGACTGAGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 300
OY 100 LeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGlyLeuAspGlu 119
DB 301 GTTCCTTGGCCGATCCGCTGGCTTATGATGATGATGATGATGATGATGATGATG 360
OY 120 GlnGlyArgLeuLeuArgArgArgArgArgArgArgArgArgArgArgArgArg 139
DB 361 CATGCTCTCTGATGAGACGACGATGATGATGATGATGATGATGATGATGATGATG 420
OY 140 LeuArgSerValSerThrAlaValLeuValLeuValLeuValLeuValLeuVal 159
DB 421 CTCGCGATGATATGCGCAGTATATAAAGCGTCTTCCAACTGATGATGATGATG 480
OY 160 AlaGlyPheMetThrProAlaGlnHisLeuGlnLeuGlyLeuSerLeu----- 176
DB 481 GTGGGTTTGTAAACGCCAACGACGACAAATTTATGAGGCAATGATGATGATGATG 540
OY 177 ProHisAsnMetPheTrpValProTyrValTrpPheAlaAsnLeuSerMetLeuValTrp 196
DB 541 AAGCAACCCAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 600
OY 197 LeuGlyGlyArgGlyLeuArgProIleLeuLeuGlnSerLeuAsnGluMetAsnThr 216
DB 601 AAGCAAGTGCATTTGGGATGATGATGATGATGATGATGATGATGATGATGATG 660
OY 217 LeuArgThrGlnCysGlyHisLeuValLeuValLeuValLeuValLeuValLeu 236
DB 661 TTCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
OY 237 ThrGlnValValLeuValAlaValLeuValSerPhePheLeuTrpCysLeuValGlyArgGln 256
DB 721 ACAAGAGTCTGACCTGGCGCTGATGATGATGATGATGATGATGATGATGATGATG 780
OY 257 PheLeuAsnProAlaValLeuValLeuValLeuValLeuValLeuValLeuVal 276
DB 781 TGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
OY 277 ThrPheLeuGlnPhePhePheTrpValGlyTrpLeuValVal 290
  
```

DB 841 ACCACTCTGAGATTTTCTTTCATGAGGCTGATTAAGTGG 882

RESULT 11

ID ABL10792/c

AB10792 standard; cDNA; 10760 BP.

26-MAR-2002 (first entry)

Drosophila melanogaster expressed polynucleotide SEQ ID NO 26858.

Drosophila: developmental biology; cell signalling; insecticide;

pharmaceutical; gene; ss.

Drosophila melanogaster.

WO2001/1042-A2.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PMD, Myers EW;

WPI, 2001-656860/75.

P-PSDB; ABB66689.

New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -

Claim 1; SEQ ID NO 26858; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of

insecticides, therapeutics and pharmaceutical drugs. The invention

discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

sequences (AB57737-AB872072).

The sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences.

Sequence 10760 BP; 3232 A; 2079 C; 2173 G; 3276 T; 0 other;

Alignment Scores:

Pred. No.:	3,556-38	Length:	10760
Score:	661.00	Matches:	169
Percent Similarity:	42.80%	Conservative:	60
Best Local Similarity:	31.59%	Mismatches:	115
Query Match:	28.13%	Indels:	191
DB:	23	Gaps:	11

US-09-622-964-5 (1-435) x ABL10792 (1-10760)

```

OY 21 LeuLeuGlySerPheGlySerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 539 CTTTACAGATGGCGAGGAGCATTTACAACTGTTGGCTAGATCTTGGCCCTTCTTG 5480
OY 41 LeuGlySerPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 59
DB 5479 ACCATTACTATGCGATCAACATGATGATGATGATGATGATGATGATGATGATG 5420
OY 59 ----- 59
  
```

Db 5419 TGAGTAATATAAGCTCCTAGAGAGATCAGTGGAGTTGGTGGCTACATTAAACATA 5360
 Oy 59 -----
 Db 5359 TTATATGAATTTGAATATATATTTTGAATTTATCTATGTCATTCATTCGAGAA 5300
 Oy 59 -----
 Db 5299 AGTGTGATGGGCTCTACTCAATACATGACATTAATTACTTAAATACACAGACAGTA 5240
 Oy 59 -----
 Db 5239 GAAAGTTTACTCAAAACATCTGAAATAGCTTTGACAACTTCACAAATGGTGTAAAT 5180
 Oy 59 -----
 Db 5179 AAAATAGCTGATGCTACAAAATACATATATATGATGACATCTTCTCCACCGA 5120
 Oy 60 -----
 Db 5119 TCTGTCTCTTTTAAAGCGCTTACTAATTAATGCAACTTCTTGCAGAACTTTGAGCC 5060
 Oy 65 LeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPheValLeuGlyPhe 84
 Db 5059 ATTGTTCAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5000
 Oy 85 TyrValThrLeuValValThrArgTyrPhePhePhePhePhePhePhePhePhePhe 104
 Db 4999 TATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4940
 Oy 105 ArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGlnGlnGlnGlnGlnGln 124
 Db 4939 CCCATCCGCGCTGTTGTCAGCTCGATCGATCGATCGATCGATCGATCGATCGATCGAT 4880
 Oy 125 ArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeuArgSerValSer 144
 Db 4879 AGCGAACAATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4820
 Oy 145 ThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAlaGlyPheMetThr 164
 Db 4819 CCGAGGGTGAAGAGGTTTCCCGGCCCAATATATATGATGATGATGATGATGATGATGAT 4760
 Oy 165 ProAlaGlnHisLysGlnLeuLysLeu-----SerLeuProHis---AsnMetPhe 181
 Db 4759 GACAATGAAGAAGCCATCATCGACATGACATGACATGACATGACATGACATGACATGACAT 4700
 Oy 182 TrpValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTrpLeuGlyLysArgIle 201
 Db 4699 TGGCTGCCCATGCTTTGGGCTCCAGTATATATACACAGGCGCCAGAAAGAGTGCAT 4640
 Oy 202 ArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGlnCys 221
 Db 4639 CCGTATGATTTTGGCTGTGAAGACCATCATGATGATGATGATGATGATGATGATGATGAT 4580
 Oy 222 GlnHisLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnValValThr 241
 Db 4579 GGACTCCCTCATCAGCTACGATACATATATGATGATGATGATGATGATGATGATGATGAT 4520
 Oy 242 ValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGln----- 256
 Db 4519 CTGGCGGTGATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 250
 Oy 257 -----
 Db 4459 GTGGTGCGCAATACACATACCTGAC----- 4424
 Oy 271 LeuValValProValPheThrPheLeuGlnPhePhePhePhePhePhePhePhePhePhe 290
 Db 4423 CTATACTTCTCTGATATATACACGCGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4364
 Oy 291 GlnLeuSerArgAlaLeuLeuGlyTyrPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 310

Db 4363 -----GCCAGTCCGTGATA-----AATCCATTGGCGAAGCATGATGAT 4322
 Oy 311 LeuGlnThrArgMetGlnCysGlnGlnArgLysValSerArgValGlnSerSerGlnAla 330
 Db 4321 TTGTAGGTACGAT-----
 Oy 331 TrpTrpArgThrProValIleProAlaThrArgGlnAlaGlnAlaGlyGlnSerLeu 4307
 Db 4306 -----AGACACATTTGTAGTCTCGAAATCA----- 4277
 Oy 351 ProGlnArgArgArgLeuTyrTrpGlnSerSerSerSerThrProLeuGlnArgMet 370
 Db 4276 -----ATAAA 4271
 Oy 371 MetIle-LeuArgProThrGlyLeuSerThrGlyIleCysArgCysProCysTrpLeuTr 390
 Db 4270 TATATCTCTCAGGTCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4211
 Oy 390 pMetArgCysThrArgThrCysLeuGlyTrpSerArgThrCysThrGly-----Ile 408
 Db 4210 GACGAGATGACCATGACATCCGAGCTTTAAAGATCAGTACGAGGAGAGTGTTC 4151
 Oy 408 rProSerHisSerProThrGlnLeuLeuProProSerSer 422
 Db 4150 CCCAAGC---AGCTGCCCTTACACAAATAGCTGCCGACATTC 4111
 RESULT 12
 ID ABL12636/c standard; cDNA; 3592 BP.
 AC ABL12636;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 32390.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EM;
 DR WPI; 2001-656860/75.
 DR P-PDB; ABB68533.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 CC
 PS Claim 1; SEQ ID NO 32390; 21pp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116175-AB120511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB157737-AB172072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIP0

ABLL12637;
26-MAR-2002 (first entry)
Drosophila melanogaster expressed polynucleotide SEQ ID NO 32393.
Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical; gene; ss.
Drosophila melanogaster.
WO200171042-A2.
27-SEP-2001.
23-MAR-2001; 2001WO-US09231.
23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
(PEKE) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EM;
WPI; 2001-656860/75.
P-PSDB; ABB68534.
New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -
Claim 1; SEQ ID NO 32393; 21pp + Sequence listing; English.

claim 1; SEQ ID NO 32393; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB101840-AB16175), expressed DNA sequences (AB101840-AB16175), expressed proteins (AB557737-AB572072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

Sequence 1345 BP; 366 A; 322 C; 323 G; 334 T. 0 other.

ment Scores:

NO.:	5.76e-32	Length:	1345
3.:	560.50	Matches:	122
ent Similarity:	49.86%	Conservative:	50
Local Similarity:	35.86%	Mismatches:	97
7 Match:	23.85%	Indels:	76
	23	Gaps:	1

-622-964-5 (1-435) x ABL12637 (1-1345), caps

QY 68 TyrCysaspserIrrYllleInLeuIleProlleSerPheValLeuClgPheYrValThr 87
 Db 7 TATGCGACAAAATGGCTCCCATTCGGTATGCTTCGTTTGGGATTTTGGTGGA 66
 QY 88 LeuValAlaThrArgTrpTrpAsnGlnTyrGluAsnLeuProTyrProAspArgLeuMet 107
 Db 67 ATCGTGTATGACCGAGATGGTGGGACCAATATACACATTCCCTGGCCAGATGAAATTGCC 126
 QY 108 SerLeuValSerGlyPheValGluGlyLysAspGluGlnGlyArgLeuLeuAlaGArgThr 127
 Db 127 ATTTTGATAGACACCGATTCATGCGTCGCATGATATAGCCAGGGATATAGGCCGAAC 186
 QY 128 LeuIleArgTyrAlaAsnLeuClgYsnValLeuIleLeuArgSerValSerThrAlaVal 147
 Db 187 ATGCTAAGATATGTGCTGCTGCCAGGATGATAGTATATTCACCAATGATATACCGCGCGT 246

[illegible]

PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EM;
DR	WPI; 2001-6556860/75.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions -
XX	
PS	Claim 1; SEQ ID NO 58; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins
CC	(sequences (AB101840-AB116175) and the encoded proteins
CC	(AB17737-AB172072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pt_sequences.
XX	

Alignment Scores:	
Pred. No.:	2,81e-28
Score:	514.00
Percent Similarity:	45.05%
Best Local Similarity:	33.9%
Query Match:	21.6%
DB:	23
Gaps:	6
Length:	2196
Matches:	123
Conservative:	105
Mismatches:	105
Indels:	95

US-09-622-964-5 (1-435) x ABL16195 (1-2196)

QY 78 IIESERPHEVALLEUGLYPHERYVALTHYLEUVALVITRAATPTIPASNGITyr 97
Db 1 ANGTCTTTGTGCTGGCGCTCTATGTGAATCTGGTGGGAACCTGGTGGAGCAATAT 60
QY 98 GLUASNIENPROTIPROASPARJLEUMETSERLEUVALSERGLYPHEVALGLUGly --- 118
Db 61 AGCGTTTTCCCTGGCCAGATACGTGGCCCTGTTCANAAAGCGGGCATATCCCACTCA 120
QY 117 -----LYASRPLUGINGLYATRGLEUENARGARGRHLEUILEATGTyAla 132
Db 121 AATGGCGGTGAATATATGAACAAGCGGTGCGCTGATGCGCTCATATATCATGCGCTATATG 180
QY 133 ASNIENGLYASNIENLEUENLEUENARGSERVALSERTHIRALAVALTyrSARPHepro 152
Db 181 GTCCTTGCGCTATGTATTAACCTCTCAGGAATTTCCCTGCTGTAAACGTCCTTTTTCCC 240
QY 153 SERALGINHISLEUVALGINALAGLYPHEMETHRPROALAGINHISLVSGLINUGlu 172
Db 241 ACAACCCAAACATCTGTGGATCGCGGTGTATGACCAATCGGAGATGAANAATCTTTCAC 300
QY 173 LYSLEUSER-----LEUPROHISASIMETRHEPTVALPROTHVALTIRPHEALAsn 190
Db 301 GCACATGANTCAGAAAGATGCCATGTCCAAATACATGATGGCTCTCGTTTGGCTACCAAC 360
QY 191 LEUSERMETHLYSALATIRPLEUGLYGLYARGILEARGSPARPROILEUENUGINSERLeu 210
Db 361 ATCATMAACCGGCGCAGAAAGATGGCGCTATGCGCTCGGATCATATGTGCAACACATA 420
QY 211 LEUASNGIUMETASNIENLEUARGTHGLINCYSGLYHISLEUITYRALATyrASPTripIle 230


```
Db 421 CTGGTAGAGCTCTCGGATATTCCTAGACGTTTAGTGGTCTTATGTGATACGACCGTC 480
QY 231 SerIleProLeuValTyrThrGlnValThValAlaValTyrSerPhePheLeuThr 250
Db 481 TGTGTCTCTTGGTCTTACACTCAGTAGTCACCTGTGTGTGCTGCTACACTATTTCATAGCT 540
QY 251 CysLeuValGlyArgGlnPheLeuAsnProAlaLysAlaTyrProGlyHisGlu---Leu 269
Db 541 GCCCTTTTGGCCGCTCAAAATGTTGCCCAACGTTCTGACAGAAAGTGACCGCAAGATCCC 600
QY 270 AspLeuValValProValPheThrPheLeuGlnPhePheThrValGlyTyrPhe-Leu 289
Db 601 GATCTGTTCTTCCCTGTTTACCGATGATGACGTTGTTTCTACGTGGCTGGCTGAG 660
QY 289 sValGlyLeuSerArgAlaLeuLeuGlyTyrPargHisGlyGlnArgGlyHisGlyGlnG 309
Db 661 G----- 661
QY 309 nLeuLeuGlnThrArgMetGlnCysGlnGlnArgLysValSerArgValGluSerSerG 329
Db 661 ----- 661
QY 329 nAlaTyrTyrPargThrProValIleProAlaThrArgGlnAlaGlnGlyGluSerLe 349
Db 661 ----- 661
QY 349 uGlnProGlyArgArgArgLeuTyrTyrGlnSerSerSerSerThrProLeuGlnArgMe 369
Db 662 -----TGGCCGAGGTGCTAATCAATCCCTTGGCGAAGAT 696
QY 369 tMetMetIleLeuArgProThrGlyLeuSerThrGlyIleCysArgCysProCysTyrPle 389
Db 697 GACGATGATATCGAGCTGAACTGAGCTAATTGACCGACACATCAAGCGCTGCTACATGATC 756
QY 389 uTyrMetArgCysThrArg-----ThrCysLeuGly----- 399
Db 757 GTGGACGAGATGCGACGAGAGACACCCCGAGCTGCTGCGGATCAGTACTGGAGTGTGTG 816
QY 400 -----TyrSerArgTyrCysThrGlyIleSerProSerH 411
Db 817 GTGCCCAAGGATCTGCCCTATAGGTGTCATCCGAACACTACCGAAGAGAGAGCCCAAG 876
QY 411 sSerProPro 414
Db 877 GGCTCCGCCG 886
```

Search completed: July 27, 2003, 11:28:06
Job time : 307.665 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 27, 2003, 11:07:41 ; Search time 3329.65 Seconds

(without alignments)
3802.121 Million cell updates/sec

Title: US-09-622-964-5
Fairfect score: 2350
Sequence: 1 MTRVTSQVANNRLGSFRL.....QLPPSSVEPPIAMPSTSA 435

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+P2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09622964/runat_22072003_101118_26013/app_query.fasta_1.2069
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=humand0.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09622964.GCN.1.17326 -runat_22072003_101118_26013 -MCPD=6 -ICPU=3
-NO_MMAR -LARGESCORE -NEG_SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2350	100.0	2420	9	AF057170	AF057170 Homo sapi
2	1925	81.9	2171	9	AF073501	AF073501 Homo sapi
3	1925	81.9	2210	9	AF057169	AF057169 Homo sapi
4	1333.5	56.7	163915	2	AC087451	AC087451 Homo sapi
5	1333.5	56.7	166867	2	AP003733	AP003733 Homo sapi
6	1302	55.4	112309	9	AC003025	AC003025 Homo sapi
7	1293	55.0	142092	9	AF139813	AF139813 Homo sapi
8	1293	55.0	196080	9	AC004228	AC004228 Homo sapi
9	1278	54.4	133683	2	AC084857	AC084857 Homo sapi
10	1252.5	53.3	160169	2	AC051664	AC051664 Homo sapi
11	1127.5	48.0	1506	9	AF440756	AF440756 Homo sapi
12	1063	45.2	1506	9	AF440758	AF440758 Homo sapi
13	1063	45.2	2500	9	AK096459	AK096459 Homo sapi
14	1060.5	45.1	1956	10	BC031186	BC031186 Homo sapi
15	1060.5	45.1	1957	10	BC031186	BC031186 Homo sapi
16	1044.5	44.4	1861	9	AK000139	AK000139 Homo sapi
17	1003	42.7	2045	9	AF440757	AF440757 Homo sapi
18	913	38.9	1289	4	AY064707	AY064707 Sus scrofa
19	801	34.1	2874	3	AF218817	AF218817 Drosophila
20	801	34.1	3610	3	AY061546	AY061546 Drosophila
21	661	28.1	32022	2	AC019521	AC019521 Drosophila
22	661	28.1	170675	3	AC008139	AC008139 Drosophila
23	661	28.1	192055	3	AC009183	AC009183 Drosophila
24	661	28.1	221888	3	AE003686	AE003686 Drosophila
25	657	28.0	147750	9	AC018761	AC018761 Homo sapi
26	644	27.4	125804	3	AC019853	AC019853 Drosophila
27	644	27.4	274351	3	AC091227	AC091227 Drosophila
28	644	27.4	163514	3	AC019853	AC019853 Drosophila
29	631	26.9	90487	9	AL592166	AL592166 Human DNA
30	600	25.5	125804	2	AC019853	AC019853 Drosophila
31	600	25.5	163514	3	AC091227	AC091227 Drosophila
32	509.5	21.7	22973	3	AE025458	AE025458 Drosophila
33	509.5	21.7	104360	10	AC124038	AC124038 Caenorhabditis
34	509.5	21.7	134412	2	AC124037	AC124037 Mus musculus
35	509.5	21.7	229161	2	AL831772	AL831772 Mus musculus
36	509.5	21.7	230447	10	AL671866	AL671866 Mouse DNA
37	504.5	21.5	181453	2	AC130166	AC130166 Rattus norvegicus
38	497	21.1	219043	2	AC073812	AC073812 Mus musculus
39	491	20.9	39590	3	CBR639K23	CBR639K23 Caenorhabditis
40	483	20.6	27394	3	CEB3286	CEB3286 Caenorhabditis
41	482	20.5	105126	2	AC014828	AC014828 Drosophila
42	482	20.5	171209	3	AC091202	AC091202 Drosophila
43	482	20.5	174551	3	AC010715	AC010715 Drosophila
44	482	20.5	174551	3	AC010715	AC010715 Drosophila
45	482	20.5	196672	3	AC004767	AC004767 Drosophila

RESULT 1

ALIGNMENTS

421 SerSerValGluProProLeuTrpAlaProProSerTrpSerAla 435
 |||||
 1365 AGTTCGTCGAGCCCTCTTATGAGGCTCCACCTTCAACATCAGCC 1409
 DB
 RESULT 2
 LOCUS AF073501
 DEFINITION Homo sapiens vitelliniform macular dystrophy protein (VMD2) mRNA,
 complete cds.
 ACCESSION AF073501
 VERSION AF073501.1 GI:3511241
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Stohr, H., Marguardt, A., Rivera, A., Cooper, P. R., Nowak, N. J.,
 Shows, T. B., Gerhard, D. S. and Weber, B. H.
 A gene map of the Best's vitelliniform macular dystrophy region in
 Chromosome 11q12-q13.1
 Genome Res. 8 (1), 48-56 (1998)
 98112782
 9445487
 TITLE
 JOURNAL JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS 2 (bases 1 to 2171)
 Marguardt, A., Stohr, H., Passmore, L., Kraemer, F., Rivera, A. and
 Weber, B. H. F.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUN-1998) Human Genetics, University, Blozentrum, Am
 Hubland, Wuerzburg 97074, Germany
 FEATURES
 source
 1. 2171
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q12-q13.1"
 /tissue_type="retinal pigment epithelium"
 1. 2171
 /gene="VMD2"
 89. 1846
 /gene="VMD2"
 /note="Best disease"
 /codon_start=1
 /product="vitelliniform macular dystrophy protein"
 /protein_id="AAC33766.1"
 /db_xref="GI:3511242"
 /translation="MTTITYSQVANAALRGSPSRLLCRGSIYLLIGETLIFILCTY
 IIRYIRALATLEEQALPEKTLICDSITLIPISVLGEFVTLVTRVMQENLPLW
 PDRMLSLVSGVEGKDEQGLRLRLTLIRYANLGNVLLRSYTAIVQRFSAQHLVOA
 GFMPBAHKDLEKLSLPHNNMFWPMVFANLSMKAWLGGRIIDPILLOSILNENKTLR
 TOCHLAAYDMSISPLVTVQVVAVVSFELTCLVGOFLNPAKAYGHEDLVVPEPE
 TFLQFFYVGLKVAEOLINFGEDDDDEFMTNVLNLOISLAVDENMODLPRMPP
 DMVKNRPEQPPVYTAASAQPRFASFMKSTFNISLNKNEDEMQPOVEDEEDACTICR
 FLGLQSDHHPFRANSTRKLLMPKRESLHGLEKKNKAQANVRGGEDKNKAMKLAV
 DAFKSFALPQPGYSAAPQTPVLPSTPMPEFLPEAPSRLKHSVGIIDRKDSLKTVSSG
 AKKSFELISEDGALEMEHPVSRKRTVEFNELDMPEIDENHLKEPLEDSPTNIHTTT
 LKIDMDMYMALENDEAHS"
 BASE COUNT 540 a 638 c 523 g 470 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.29e-112 Length: 2171
 Score: 1925.00 Matches: 366
 Percent Similarity: 84.17% Conservative: 1
 Best Local Similarity: 83.94% Mismatches: 0
 Query Match: 81.91% Indels: 69
 DB: 9 Gaps: 1
 US-09-622-964-5 (1-435) x AF073501 (1-2171)

QY	21	LeuLeuCystrPrArgGlySerIleYrrYrsLeuLeuYrGlyGluPheLeuIlePheLeu	40
Db	149	CTGGCTGCTCGCGGGGAGCATCTACAGCGGCTATATGGGAGTTCTTAACTCTCTG	208
QY	41	LeuCystrYrrYrIleIleArgPheIleYrrArgLeuAlaIleThrGluGlnGlnLeu	60
Db	209	CTCTGCTACTACATATATCGGCTTTTATTAGCGTGGCCCTTCACGGAAACACACTG	268
QY	61	MetPheGluYrsLeuThrLeuYrCysAspSerYrIleGlnLeuIleProIleSerPhe	80
Db	269	ATGTTTGAGAAACSTGACTCTGTAATGGACAGCTACATCCAGCTCATGCCATTTCTTC	328
QY	81	ValLeuGluYrPheYrValThrLeuValValThrArgYrrPrPArgGlnYrGluAsnLeu	100
Db	329	GTGGCTGGGCTTCTAGTAGCACTGGCTCTGACCCGCTGGTGGAACCATAGACAGAACCTG	388
QY	101	ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAspGluGln	120
Db	389	CCGTGGCCCGACCGGCTCATAGACCTGGTGTCCGGCTTCGTTCGAAGGCACAGCAGCA	448
QY	121	GlyArgLeuLeuArgYrArgThrLeuIleArgYrAlaAsnLeuGlyAsnValLeuIleLeu	140
Db	449	GGCCCGGCTGCGCGGCGCACGCTCATCCGCTAACGCCAACCTGGGCACAGTGCCTCATCTG	508
QY	141	ArgSerValSerThrAlaValThrYrLysArgPheProSerAlaGlnHisLeuValGlnAla	160
Db	509	CGACGCGTACACACCGCATGCTACACAGGCTTCCCGACGCCACACCTGGTGGACCA	568
QY	161	GlyPheMetThrProAlaGluHisLysGlnLeuGluYrsLeuSerLeuProHisAsnMet	180
Db	569	GGCTTTATGACTCCGGCGAACAACAAGCAGTTGGAGAACTGAGGCTTCCACACACACTG	628
QY	181	PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyGlyArg	200
Db	629	TTCTGGGCGCCCTGGGCTGTGGTTGGCCAACTGTCAAGAGCGGTGGAGTGTGCA	688
QY	201	IleArgAspProIleLeuGlnSerLeuLeuAsnGluMetAlaThrThrLeuAlaThrGln	220
Db	669	ATCCGGAGCCCTATACCTCTCCACAGGCTGTGAAACGATGAACACCTTCGTACTAG	748
QY	221	CysGlnHisLeuYrAlaYrAspTrpIleSerIleProLeuValYrThrGlnAlaVal	240
Db	749	TGTGGACACCTGTAATGCTTACGACTGATTAATCCACAGGTGTATACACAGTGGTG	808
QY	241	ThrValAlaValYrYrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro	260
Db	809	ACGTATGGGCGTGAACGCTTCTTCCTGACTGTCTAGTGGCGGACATTCTGAACCA	868
QY	261	AlaIysAlaYrTrpGluYrHisGlnLeuAspLeuValValProValPheThrPheLeuGln	280
Db	869	GCCAAAGGCTAACCTGGCCATGACCTGGACCTCTGTGGCCGCTTCACACTTCCTGAG	928
QY	281	PhePhePheYrValGlyTrpLeu-LysValGlyLeuSerArgAlaLeuLeuGlyTrpArg	300
Db	929	TTCTCTTCTCTATGTTGGCTGGCTGAGG-----	956
QY	300	GHisGlyLeuArgGlyHisGlyGlnGlnLeuLeuGluThrArgMetGlnCysGlnGluArg	320
Db	956	-----	956
QY	320	GlyValSerArgValGluSerSerGlnAlaTrpTrpArgThrProValIleProAlaThr	340
Db	956	-----	956
QY	340	ArgGluAlaGluAlaGlyIuSerLeuGluProGlyArgArgArgLeuTrpTrpGlnSer	360
Db	957	-----TGGCAGAG	964
QY	360	rSerSerSerThrProLeuGluArgMetMetMetIleLeuArgProThrGlyLeuSerThr	380
Db	965	CAGCTCATCAACCCCTTGGAGAGATGATGATGATTTTGAGACCAACTGATGTGTGCAG	1024

	QY	380	rglyllcCysArGdySProcSTPleATrPMeAryCysThrArgThrcysLeuGlyTr	400
	Dd	1025	AGCAATTTCGACGTCTCCCTGTTGGCTGTGCATGATGCACACGACTCGCTCGCATG	1084
	QY	400	pSerArgThrCysThrGlylleSerProSerHisSerProProthrGlnIleuLeuProPr	420
	Dd	1085	GAGCGGACACATGTACTCGAATAAGCCGAGCACAGGCCCCCTTACACAGCGCTCCGCC	1144
	OY	420	oSerValGluProProLeuTriPalalProProSerThrSeraL	435
	Dd	1145	CAGTTCGGTCGAGCGCTCTTATGGGCTCCACACTTACAACATCACC	1190
RESULT 3				
AF057169				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
SOURCE				
gene				
CDS				
BASE COUNT				
ORIGIN				
Alignment Scores:				
Pred. No.:				
Score:				
Percent Similarity:				
Best Local Similarity:				

Query Match:	81.91%	Indels:	69
DB:	9	Gaps:	1
US-09-622-964-5. (1-435) x AF057169 (1-2210)			
QY 1 MethrllethrrYrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20			
DB 105 ATGACACATCACTTACACACAGCCAGCTTAATGCCCGCTTAGGCTCTCTCCGCCGCTG 164			
QY 21 LeuLeuCYSTPAArgGlySerLleuTrLysLeuLeuTrGlyGluPheLeuIlePheLeu 40			
DB 165 CTGCTGTCTCTGGCGGGGAGCATCAACAAGCTGCTATATGGGAGATTCTTAATCTTCCG 224			
QY 41 LeucystYrTrIleIleArgPheIleYrArgLeuAlaLeuThrGluGluGlnGlnLeu 60			
DB 225 CTGTCGCTACTACATCATCATCCGCTTATTATATAGCGTGGCCCTCACGGAAAGAACAGCTG 284			
QY 61 MethheGluLysLeuThrLeuTrLysAspSerTrIleGlnLeuIleProLieserPhe 80			
DB 285 ATGTTTGGAAGAACCTGACTCTGATTGGACACGCTACATCCACTCATCCCTATTCCTCTG 344			
QY 81 ValIeuGlyPheYrValThrLeuValValThrArgTrPTrAsnGlnTrGlyLysAsnLeu 100			
DB 345 GTGCTGGGCTTCTACGTACACCTGGTCTGTGACCCGCTGGTGGAACACAGTACGAAACCTG 404			
QY 101 ProTrPProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120			
DB 405 CCGTGGCCGACCGGCTCATGAGCTGGTGTGGCTTCTGTGAAAGGCAAGGCGAGCA 464			
QY 121 GlyArgLeuLeuArgArgThrLeuIleArgTrAlaAsnLeuGlyAsnValLeuIleLeu 140			
DB 465 GGGCGGCTGTGGGGCGACGCTCATCCGCTACGGCAACCTGGGCAACCTGCTCATCTCTG 524			
QY 141 ArgSerValSerThrAlaValTrLysArgPheProSerAlaGlnHisLeuValGlnAla 160			
DB 525 CGCACCGCTACGACCGGCACTGTACAAAGGCTTCCCGAGGCCACAGCACTGGTGCAAGCA 584			
QY 161 GlyPheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet 180			
DB 585 GGGCTTTATGACTCCGCGAAGACACAAAGGTTGGAAACTGAGCCTTACCACCAACATG 644			
QY 181 PheTrpValProTrpValTrpPheHisAsnLeuSerMetLysAlaTrPheGluGlyArg 200			
DB 645 TTCGTGGGGCCCTGGGTGGTTCGCCAACCTGTCAATAGAGCGTGGCTTGGAGGTCCA 704			
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220			
DB 705 ATCCGGGACCCATATCTGCTCCAGAGCCGCGTGAACGATGAACACCTTGGCTACTCAG 764			
QY 221 CysGlyHisLeuTrAlaTrAspTrPTrIleSerIleProLeuValTrpThrGlnValVal 240			
DB 765 TGTGGACACCTGTATGCTTACACACTGAGTATATCCACTGGTGTATACACAGGGGTG 824			
QY 241 ThrValAlaValTrpSerPhePheLeuTrGlySLeuValGlyArgGlnPheLeuAsnPro 260			
DB 825 ACTGTGGGGGTATAAGCTTCTTCCGACTTGTCTTAAGTTGGGGGCAAGTTTCGAAACCA 884			
QY 261 AlaLysAlaTrpProGlyHisGlnLeuAspLeuValValProValPheThrPheLeuGln 280			
DB 885 GCGAAGGCTTACCCGAGCATGAGCGGACGTCGTGTGCCCGCTTACAGTTCCTGCAG 944			
QY 281 PhePhePheTrpValGlyTrpLeu-LysValGlyLeuSerArgAlaLeuLeuGlyTrpPar 300			
DB 945 TTCCTCTCTATAGTGGTGGCTGTAAGG----- 972			
QY 300 GHisLysGlnArgGlyHisGlyGlnLeuLeuGluTrpTrpArgMetGlnCysGlnGlu 320			
DB 972 ----- 972			
QY 320 GlyValSerArgValGluSerSerGlnAlaTrpTrpArgThrProValIleProAlaThr 340			
DB 972 ----- 972			

QY		340	rargGuaIagIuaIacgysusertLeuglProglYargrayrgrLeutrPrpJnse	360
Db		973	-----TGGCAGAG	980
QY		360	rSerSerSerThrProleugluArgMetMeUetlleUauRgrProthrglyLeuSerTh	380
Db		981	CAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGAGACCACACTGGATTGTGCAC	1040
QY		380	rglyllcgcYargcysProCystTrPleUrTmeArGcysThrArghrCysleuGlyTr	400
Db		1041	AGGAATTTCGACGGGTGCCCTGGTGGCTGGATAGANGCACCGAGACTCCCTCGGATG	1100
QY		400	pSerArgThrCysThrGlyIleSerProSerHisSerProprorhrgInleuLaupProp	420
Db		1101	GAGCCGGACATGTACTGTGATAAAGCCGACGCCAACACCCCCCTTACACAGCTTCGCC	1160
QY		420	oSerServAlgluProProleUrtpalaProProserThrsEraLa	435
Db		1161	CAGTTCGGTCGAGCCTCCTTATATGGGCTGCACCTTCACATCAGCAGC	1206
RESULT 4				
LOCUS	AC087451/c	163915 bp	DNA	linear HTG_04-JUL-2001
DEFINITION	Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT SEQUENCE, 7 unordered pieces.			
ACCESSION	AC087451			
VERSION	AC087451.2 GI:14595836			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLPROP.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.			
TITLE	1 (bases 1 to 163915)			
JOURNAL	Homo sapiens chromosome 11, clone RP11-810P12			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 163915)			
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Batte... Batna,N., Batten,V., Boguslavsky,I., Boukhgalter,B., Brown,A., Camarile,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S., Collumock,A., Cooke,P., Dellrelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,U., Hulme,W., Illey,T., Johnson,R., Jones,C., Karatas,A., Larocque,K., Lamazaras,R., Landers,T., Lehocskyy,J., Levine,R., Liu,G., Maclean,C., MacDonald,P., Marcynski,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McNeesters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenaga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,F., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhuang,P., Plerie,N., Pollara,V., Raymond,C., Retta,R., Ribback,M., Riley,R., Riser,C., Rogov,P.P., Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schnuppack,R., Seaman,S., Severly,P., Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausz,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M., Travls,N., Triggillo,U., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
TITLE	Submitted (05-JUN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
JOURNAL	On Jul 4, 2001 this sequence version replaced gi:12039464.			
COMMENT	All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html			
	----- Genome Center			
	Center: Whitehead Institute/ MIT Center for Genome Research			
	Center code: WIRB			
	Web site: http://www-seq.wi.mit.edu			
	Contact: sequence.submissions@genome.wi.mit.edu			
	----- Project Information			
	Center project name: L11863			
	Center clone name: 810_P_12			

```

----- Summary Statistics -----
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.96071
Consensus quality: 162458 bases at least Q40
Consensus quality: 162916 bases at least Q30
Consensus quality: 163173 bases at least Q20
Insert size: 170000; agarose-ftp
Insert size: 163315; sum-of-ctrlgs
Quality coverage: 10.5 in Q20 bases; agarose-ftp
Quality coverage: 10.9 in Q20 bases; sum-of-ctrlgs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 17285: contlg of 17285 bp in length
* 17286 17385: gap of 100 bp
* 17386 19170: contlg of 1785 bp in length
* 19171 19270: gap of 100 bp
* 19271 28600: contlg of 9330 bp in length
* 28601 28700: gap of 100 bp
* 28701 47671: contlg of 18971 bp in length
* 47672 47771: gap of 100 bp
* 47772 105063: contlg of 57292 bp in length
* 105064 105163: gap of 100 bp
* 105164 142181: contlg of 37018 bp in length
* 142182 142281: gap of 100 bp
* 142282 163915: contlg of 21634 bp in length.
*
Location/Qualifiers
1. 163915
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone_RP11="RP11-810P12"
/clone_lib="RP11 Human Male BAC"
1. 17285
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
17386. 19170
/note="assembly-fragment"
19271. 28600
/note="assembly-fragment"
28701. 47671
/note="assembly-fragment"
47772. 105063
/note="assembly-fragment"
105164. 142181
/note="assembly-fragment"
142282. 163915
/note="assembly-fragment"
clone_end:T7
vector_side:right"
BASE COUNT 43950 a 38090 c 37945 g 43329 t 601 others
ORIGIN
Alignment Scores:
Pred. No.: 1,32e-72 Length: 163915
Score: 1333.50 Matches: 383
Percent Similarity: 23.24% Conservative: 0
Best Local Similarity: 23.24% Mismatches: 1
Query Match: 56.74% Indels: 1264
DB: 2 Gaps: 6
US-09-622-964-5 (1-435) x AC087451 (1-163915)
51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluGlyLeuThrLeuTyrCysAsp 700

```

Db 158577 AGGCTGGCCCTCAGGAGAAACACAGCTGATGTTGAGAAACTGATGATTCGAC 158518
QY 71 SetTyIleuIleuIleuProIleuSerPheValIleu----- 82
Db 158517 AGCTACATCCAGCTCATCCCATTTCTCTGCTGGGTAGTTCCTCCCTTCGCGCTT 158458
QY 82 ----- 82
Db 158457 CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGCCAGAGGAGATCAGAGAGCTCG 158398
QY 82 ----- 82
Db 158397 GCAAGGGGCTGGGAGGGGGGGGGAACCCAGCGGAGGTCGGCGCTCTGTAGGG 158338
QY 82 ----- 82
Db 158337 AAGGTCCGACCTGCAGCCAGTGAACCTGAAGTTAGCTTAGAGAGCTCCTGCCGT 158278
QY 82 ----- 82
Db 158277 TAGCAATGAACCCCATTTTCTGAGGGAAGCGCTGACATCATGCTCCCTGAGCCCTG 158218
QY 82 ----- 82
Db 158217 CGCGGAGGGAGGGGGTCTGGCGAATTCTGGGACAGCAGGGGGACCCCGGGTGACA 158158
QY 82 ----- 82
Db 158157 GAACCTTGGGGCTCTGGCCCTCCATGCGAGCTCTGCCCTCTCGCTCCGAGGCG 158098
QY 82 ----- 82
Db 158097 CTTCAGAGAGGCTGGGGGCTAGGCCCTCGCAGCAGAAAGCTGAGAGCCGAGCAT 158038
QY 82 ----- 82
Db 158037 CGCGGGGCGTGGCCCTGGGCTGCGCCGACGCTGGCCCTGCCCTGCGCCCGC 157978
QY 83 -----glyPheTyValThrLeuValValThrArgTPTPasnglnTygIu 98
Db 157977 CCCTCTGCCCCAGCTTCTAGCTGAGCGCTGCTGCGAGCCGCTGGTGAACAGTACGAG 157918
QY 99 AsnLeuProTPProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyAsp 118
Db 157917 AACCTGCGTGGCCCGACCGCTCATGAGCTGGGTGCTGCTGGAAGGCAAGGAC 157858
QY 119 GlnGlnGlyArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
Db 157857 GAGCAAGGCCGCGTCTGCTGGCGGACGCTCATCCGCTACGCCCAACCTGGGCAAGCTGCTC 157798
QY 139 IleLeuArgSerValSerThrAlaValTyrIlySarGpheProSerAlaGlnHisLeuVal 158
Db 157797 ATCTGGGCGAGGTGAGCAGCAGCACTACAAAGCCTTCCGAGCGCCAGCACTGGTG 157738
QY 159 Gln----- 159
Db 157737 CAAAGAGTGGGCGGAGCCGAGCAACGGGAGGAGCCGGGAGAGCCAGAGGGCCGAGAT 157678
QY 159 ----- 159
Db 157677 GGGCGCGCAGGAATGGAAGATGGGTGAGGCAAGTCCCGCGAGCTGGGGGATTTGGGT 157618
QY 159 ----- 159
Db 157617 GGAGCAGAGAGTGGGGTGTAGTCAAGATTTGGGGTCCAAATTGGCGGGAGACAGTCCGG 157558
QY 159 ----- 159
Db 157557 TGTCTGAAGGTGGGCGAGGCCAGGACCCACTCCGAGAGTAGAGTCTGAGGCGAGG 157498
QY 159 ----- 159

Db 157497 CTAGAGACCCCTTGAGGGATATGGAAGAGGGTGAAGGCTTGGAACTGTGAGGTACT 157438
QY 159 ----- 159
Db 157437 AGGCTACTTCCCTCTGCCCTTGGCCCTCTTGATGATCCGGTTCCACTCTGAGGTATG 157378
QY 159 ----- 159
Db 157377 GGACATTGCTCTGACACCCCTCAAGCCCTGAGCTGACCTGTCGTGGTTAATAAGACAG 157318
QY 159 ----- 159
Db 157317 ACCCAGGCTAGGGGTGGGTGCTGCTCGCTGTAATCCAGTCTTTAGAGGCAAGGTGG 157258
QY 159 ----- 159
Db 157257 GAAGATCGCTTGAAGCCAGCTGTTAGAGACGCCCTGAGCAACATAGCGAGACCCCATC 157198
QY 159 ----- 159
Db 157197 TCTACAAAACATTAAAAATTACAGGGCATGGTGGCGTGGCTGAGTGTAGGCTGA 157138
QY 159 ----- 159
Db 157137 GTATCGGGAGCTGAGGCGAGGAGATCATTGAGCCAGAGTTCCAGGCTGACGTGCGC 157078
QY 159 ----- 159
Db 157077 TAAGATCGACCGCTGCACATCCAACTGCTGAGACAGACCACTTTCTCTGGAATA 157018
QY 159 ----- 159
Db 157017 AATTAATACCTGCCCCACATGCTCAGCCAGAACAGCACCATGTAAGTGTCTAGAAATTT 156958
QY 159 ----- 159
Db 156957 TTTGTGTGAAGAAAGAGATGCGAAAGAGTGTGAGTTCTATAGTACGAGG 156898
QY 160 -----AlaG1 161
Db 156897 TGCGGGCCATCCCTTCTGACAGTTCTCCACCCACCGCTTCTTCACTCCACTGCGAC 156838
QY 161 yPheMetThrProAlaGlnHisIlysglnLeuGlnIlyLeuSerLeuProHisAsnMetCph 181
Db 156837 CTTTATGACTCCGGCAGAACACAGCAAGTGTGAGAACTGAGCTTACCAACACATGTT 156778
QY 181 etrPValProTPValTirPheAlaAsnLeuSerMetIlySarTirPheGlnIlyArgII 201
Db 156777 CTGGGTGCCCTGGGTGTGTGTGGCCAACTGTCAATGAAGCGGTGGAGGTGCAAT 156718
QY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn----- 212
Db 156717 CCGGAGCCCTATCTGCTCCAGAGCTCTGTAACTGAGCCCACTGTACAGACAGGGCTG 156658
QY 212 ----- 212
Db 156657 CCGCAGATGGGAAGGCTGTGTCTCAGAGAAACAAAGTTCTTACAAGAAAGCCTT 156598
QY 212 ----- 212
Db 156597 GGGCCCTGAGAGGTCTTCCAGAGCTGAGGTGGGGTTGCAAGATCTTTTCCACAGCAA 156538
QY 212 ----- 212
Db 156537 TCCACAGCCCGAGGTGCTCCTTCCAGAGGCCCTCCCTTCTTCCAAGTGTGAGGT 156478
QY 212 ----- 212
Db 156477 CCTGTTCCCTTTGTATGATGAGGAGACTGAGACACAAGAGTTTATGAGCTTCCA 156418
QY 212 ----- 212
Db 156417 TGGCCACAGCCAGGAATGACCATAGGTACAGGCCCTGTGTTACCTGAGAGAAGGTGG 156358

QY	212	-----	212	-----	212
Db	156357	GGGCGACCCAGGCTGGGGGCGAGGTGGTTCAGAAACCCATCCCTCTTCGCCCC	156298		
QY	213	--gluMetAsnThrLeuArgThcIncysglYhIsleuTYrAlaTYrAspTrpIleSeri	232		
Db	156297	AGGAAATGAACACCTTCGTAATCAGTGTGGACACCTGATAGCTTACGACTGGATTAGTA	156238		
QY	232	leuProleuValTYrThr	237		
Db	156237	TCCCACTGGTGTATACACAGGTGAGACTAGAGCTGGTGGAGGCTGCCCTTTGGGAAACTG	156178		
QY	237	-----	237		
Db	156177	AGGCTAGAAAGACCAAGGAAGACAGCTGGGGTGGGAAGGGCTCACCTAGAGGCTAAAGTGC	156118		
QY	237	-----	237		
Db	156117	TCCCCTGGAGTTGGGTCCACACTTTGAAGTTGGGTCTGGACTTTGAAGTCCAAAGTTCT	156058		
QY	237	-----	237		
Db	156057	AAGAGTCCAGGCTCCTGCCTGGCCACGTCAGTAGAGCAATGTGATTAATCCCATATTA	155998		
QY	237	-----	237		
Db	155997	AAGAGAGGTGGCGCGGGTGCAGGTGCTCATGCTGTAAATCCACAGCACTTTGGGAAGTGA	155938		
QY	237	-----	237		
Db	155937	GGCAGGTGATACCTGAGAGTCCAGAGTTCCAGACACAGCTGGCCAAATGATGAAACC	155878		
QY	237	-----	237		
Db	155877	CATCTCTACTGAATAATACAGAAATAGCTGTGTGGTGCACGCTGTAAATCCACAGCTAC	155818		
QY	237	-----	237		
Db	155817	TTGGAGAGCTGAGCGACGAGAAATCGCTTGAACCCGGAGAGTGGAGTTGCAGTGAGTGA	155758		
QY	237	-----	237		
Db	155757	GATCATGCGCACTGCACTCCAGCGCTGGCGACACAGCAAGACTGTCTCAACAAACAA	155698		
QY	237	-----	237		
Db	155697	CAAAACAACAAACAAAGGGTTAACAGAGCCCTTAATGTCACATAGTGTCAAATGCAGA	155638		
QY	237	-----	237		
Db	155637	CAAGGCTTGCTGCTCTGTCTGACTCCAGCCCTGGAGCATCTGATTCAGGGTTC	155578		
QY	238	-----	245		
Db	155577	CCACCTAGCCCTTGTGTACACATCTCTCTCTCCACAGGGTGGAGCTGGGGCTGT	155518		
QY	245	YrSerPhePheLeuThrCysLeuValAlYArgInPheLeuAsnProAlaYsAlaTYrP	265		
Db	155517	ACAGGCTCTCTGACTGTCTGTCTGTGGGGGAGATTTCGAACCCAGCAAGGCTTACC	155458		
QY	265	roglYhIsGluLeuAspLeuValValProValPheThrPheLeuGlnPhePhePheTYrY	285		
Db	155457	CTGGCCATGAGCTGAGACCTCGTGTGGCCGCTTTCACGTTCGACAGTCTCTTCTTATG	155398		
QY	285	alglYTrPLeuYsValGlYLeuSerArgAlaLeuLeuGlYTrPArgHIsGlYAlArgG	305		
Db	155397	TTGGCTGGCTGAAGGTGGGCTCTCCAGGGCCCTGCTGGGCTGAGAGCATGGCCAAAGGG	155338		
QY	305	LYhIsGlYGlNglNLeuLeuGlUthrArgMetGlnCysGlnGluArgYsValSerArgY	325		
Db	155337	GTCATGGCCACAGACTCCTCGAAGACGAGATGAGGTCTCAGAAAGGAAGTCTCACGGG	155278		

QY	325	algiuserSerGlnAlaTrpTrpArgThProValIleProAlaThrArgGluAlaGln	345
Db	155277	TAGAAGCAGCAGCAGCGCTGGTGGCGCACACCTGTAAACCCAGCTACTCGGGAGGCTGAGG	155218
QY	345	laagiuserLeuGluProGlyValArgTrpArgLeuTrp	357
Db	155217	CAGAGAAATCGCTTGAACCCGGGAGGCGAGGTTGTGTGATGTTGAGATGCTGCACCTGC	155158
QY	357	-----	357
Db	155157	ACTCAGCCTGGGCAAAAGATGAAGACTATCTCAAAAACACACACAAACAAACAA	155098
QY	357	-----	357
Db	155097	AGCCCTAAGTTCAGAAAGCCCTCGCCTTTAGAAAGCAGATCGGCACACCTCTCTTAT	155038
QY	357	-----	357
Db	155037	TCAGATGCTGTGGGCTGTCTTGTTCCTCACTCAAGTGGCTTGTCAAGTATTCCT	154978
QY	357	-----	357
Db	154977	CCACACAGCAGCAATACTCCGAAACAGATGTCTGAATCACACAGTTTCTCTCACTCT	154918
QY	357	-----	357
Db	154917	TTATCTTCCCTCTCTGTGTGTGCCACACCACTCTCTCCCTCTCACTCTTATTT	154858
QY	357	-----	357
Db	154857	TTTGGTAATGGGGGTGAAGTCTGTGTCTGTCCCTCTCTGACTGTGACACACACACA	154798
QY	357	-----	357
Db	154797	CACACACACACACACACACATACACACACACACACAGCATTCCTATTCCTAAATT	154738
QY	357	-----	357
Db	154737	CCCCCTGCCCCCCAGTTATCTTTGGTTTCTGCAGATCAAAACAATCACTTTATGC	154678
QY	357	-----	357
Db	154677	TTGAAATTCTCCGGGGTCCCCCAGTGGCTGCAAGATGTCCCTGGACCCCTTAAGCAGA	154618
QY	357	-----	357
Db	154617	CGGTGTCAACCTTTCGGGGCTTTGTTAGGGCATTTTAGAGTTGCTATCAGAAATCTG	154558
QY	357	-----	357
Db	154557	CCCACCTAGACTGCCCTTAGTTAGTTCAGCCAGCTTCAGTATATATCTCTGTGCATGAATG	154498
QY	357	-----	357
Db	154497	AATATAATTATGCACTCCAGTAAAGATATACATGAGTGAGATTAAGCAGTACTCAGCC	154438
QY	357	-----	357
Db	154437	GAGTATACACTCAGGACAGCTGTGGGTGTTCAGGAAAGCACTGCTCAGAAAGCTTAG	154378
QY	357	-----	357
Db	154377	AGGGGTGTGTCCAGAAATGTGTGGGTGCCCAAAAGTGTGGGGGGCTGGAGCCCTAACT	154318
QY	357	-----	357
Db	154317	CTGCTTTGAAGACAGTGTCTCAGGACGAAGGCGTATGGGTGTGAATATGACAGCAG	154258
QY	357	-----	357
Db	154257	CTGAGTTTAAAGGGGAGCTGTGCTTTGAGAGATTCTGCTTGAAGGTTTACAGACCTTC	154198
QY	358	-----TtgcIuserSerSerSerThrProLeuGluArgMetMetMetIleIle	373

Db	Accession	Gene	Species	Accession	Gene	Species	Accession	Gene	Species
Db	154197	ACGCTGCCCAAGGTCGACGACGACCTCATCAACCCCTTTGGAGAGGATGATGATGATTT	154138						
Oy	373	euatgprctthgyleuserthrglyilecys	383						
Db	154137	TGAGACCAACTGGATTGTTCGACAGGAATTTGCAGGTATGGGAGAGAGGAAACCATTA	154078						
Oy	383		383						
Db	154077	CCATGGACCTTCCCAAGATGGACCCCAAGAGAGACCCCACTGTTCTGTAGGAGGCCCT	154018						
Oy	383		383						
Db	154017	CACAGTGAATGATCAACCTTCCCTCCTCCTGTGACGACGATCATTCATCACAGGA	153958						
Oy	383		383						
Db	153957	TTTCACCTCAATCTTTGAGGCTGACAGGACGCCACCCATCTCCCATTTACAGGACGG	153898						
Oy	383		383						
Db	153897	AAACTGAGTCCAGAGAGAGGAGAGATTCCTCAATCATGACGACATACAAAGTCTT	153838						
Oy	384		388						
Db	153837	GCCCTGGATGATCTTCTGTGGACTTCTTCTGCTCCCTGGTGCACAGGCTCCCTGTGG	153778						
Oy	389	LeuTPmetarGcysThrArgrthrcysleuGlyTrpSerArgrthrcysThrglyIleSer	408						
Db	153777	CTGTGGATGAGATGCACGACGACGACGCTCCGATGGAGCGGACGACATGTGAAATAGC	153718						
Oy	409	ProSerHisSerProProThrGlnLeuLeuProProSerSerValGlnProPoleuTrp	428						
Db	153717	CCGACCCACAGCCCCCTCACAGAGCTTCCGCCACGATTCCTCGAGCCTCTTATGG	153656						
Oy	429	AlaProProSerThrSer	434						
Db	153657	GCTCCACCTTCAACATCA	153640						
RESULT 5	AP003733	166867 bp	DNA	linear	PRI 08-DEC-2001				
LOCUS	AP003733	166867 bp	DNA	linear	PRI 08-DEC-2001				
DEFINITION	166867 bp	DNA	linear	PRI 08-DEC-2001					
ACCESSION	AP003733	166867 bp	DNA	linear	PRI 08-DEC-2001				
VERSION	AP003733.4	GI:17426128							
KEYWORDS	HTG.								
SOURCE	Human sapiens DNA, clone:RP11-810P12.								
ORGANISM	Human sapiens								
REFERENCE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
AUTHORS	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
TITLE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
JOURNAL	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
REFERENCE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
AUTHORS	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
TITLE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
JOURNAL	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
REFERENCE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
AUTHORS	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
TITLE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
JOURNAL	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
REFERENCE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
AUTHORS	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
TITLE	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totsuki,Y., Watanabe,H. and Sakaki,Y.							
JOURNAL	1	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada							

ORIGIN		BASE COUNT	44705 a	38465 c	39268 g	44429 t
--------	--	------------	---------	---------	---------	---------

Alignment Scores:						
Pred. No.:	1,34e-72	Length:	166867			
Score:	1333.50	Matches:	383			
Percent Similarity:	23.24%	Conservative:	0			
Best Local Similarity:	23.24%	Mismatches:	1264			
Query Match:	56.74%	Indels:	1			
DB:	9	Gaps:	6			

US-09-622-964-5 (1-435) x AP003733 (1-166867)						
---	--	--	--	--	--	--

QY	51	ArgLeuAlaLeuThrGluGlnGlnGlnIleMetPheGluLysLeuThrLeuTyrCysAsp	70
Db	5339	AGCGTTGGCCCTCACACGGAAGAACAACACTCATGTATTAGAAACAGACTCTGTATGGAC	5398
QY	71	SerTyrIleGlnLeuIleProIleSerPheValLeu-----	82
Db	5399	AGCTAACATCCACAGCTCATCCCATTCCTTCGTGGTGAGTTCCTCCCTTCGTGCTGT	5458
QY	82	-----	82
Db	5459	CCGGGTCCCTGTGGCCGCCACAGGCTCAGACAGGCCAGGGAGATACAGAGACTGCG	5518
QY	82	-----	82
Db	5519	GCAAGGGCTGGGAGGGGGGGGGAACCGCACGGGCAGGTGGCGCCTCTGTAGGG	5578
QY	82	-----	82
Db	5579	AAAGTGGCGAGCTGCAGCCACTGAACGTGAAGTTAGAGTTAGGTAAAGAGTCTCGCT	5638
QY	82	-----	82
Db	5639	TAGCAATAAACCACCATTTTCTGAGGGAAGCGCTGACATCATGTGCCCTGGAGCCCTG	5698
QY	82	-----	82
Db	5699	CGCGGAGGGAGGGGGGCTGTGCGGATTTCTGGAGCACAGAGGGGACCCCCGGGTGACA	5758
QY	82	-----	82
Db	5759	GAACCTTGGGGCTCTCGCGCTCCATCGAGGCTGTGCTCTCTCGTCCGAGCGC	5818
QY	82	-----	82
Db	5819	CTTCAGAGAGGGCTGGGGGCTAGGGCCCGCTGCAGCAGAAAGGTGAGAGCGAGGCAT	5878
QY	82	-----	82
Db	5879	CGCGGGGCGTGGCGCTGGGCTGTGCGCGAGCTGTGGCCCTCGCCCTCGCCCGCCG	5938
QY	83	-----GlyPheTyrValThrLeuValValThrArgTPTPrAsnGlnTyrGlu	98
Db	5939	CCCTCTGCGCCAGCGCTTCTACGTGACGTGGTCGGAGCCCGCTGTGGAAACCACTACGAG	5998
QY	99	AsnLeuProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp	118
Db	5999	AACCTGCGCGTGGCCGACCGCTCTAGAGCGTGGTGGGCTTCGTGGAAGCAAAGGAC	6058
QY	119	GluGlnGlyArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu	138
Db	6059	GAGCAAGGCGCGCTCTGTGGGGCGACGCTCATTCCTTCAGCCAACTGGGAACGTGCTC	6118
QY	139	IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal	158
Db	6119	ATCTGTGCGAGGTAGCAGCCGCACTCTACAGCGCTTCCCAGCGCCACACTGTGTC	6178
QY	159	Gln-----	159
Db	6179	CAGAAGGTGGGCGGAGCGGAGGCAACGGGGGAGGACCGGGGAGCGAGGGGCCGAGAT	6238

159 ----- 159
Db 6239 GGGCGCGCAGAGATGGAAGATGGGTGGAGCCAAAGTCCCGGAGCTGGGGATGGGT 6238
159 ----- 159
Db 6239 GGAACGAGAGTGGGTGTAGTCAAGATTTGGGGTCCAAATTTGGGGGACAGAGTGGG 6238
159 ----- 159
Db 6359 TGTCTGAAGTGGGGCGAGGCGAGGCCACCTCCGAGAGTAGAGTCTGAGGCGAGG 6418
159 ----- 159
Db 6419 CTAAAGACCTTGAAGGATATGGAAGAAAGAGGTGACGGCTTGGGAACTGGTAGTACT 6478
159 ----- 159
Db 6479 AGGGCTACTTCCCTCTGCCCCCTTCTGATCTCCGGTTTCCACTCTGGAGGTATG 6538
159 ----- 159
Db 6539 GGACATTGGTCTGTGACACCCCCCTCAGCCTGGCTGACCTGGTCTGTATATAGACAG 6598
159 ----- 159
Db 6599 ACCGAGCTAGGCGTGTGCTGTGCGCCTGTAAATCCAGTGTCTTGAAGAGCAAGTGG 6658
159 ----- 159
Db 6659 GAAAGTGGCTTGAGCCAGCTGTGTTGAGACGCCCTTGAGCAACATAGGAGACCCCCATC 6718
159 ----- 159
Db 6719 TCTACAAAACATTAATAATAGCAGGCGATGGTGGCTGTGCTGCACTGAGGCTGA 6778
159 ----- 159
Db 6779 GTATCGGAGGCTGAGGCGAGGAGATCACTTGAGCCAGAGTCCAGGCTGAGTGGC 6838
159 ----- 159
Db 6839 TAAGTCGACCGCTGCACTCCACCTGGTGACAGAGCAGACCTTTCTGTGAAATA 6898
159 ----- 159
Db 6899 AATAAATACCTGCCACATGTCTAGCCAGAACAGACACTAGTAGTGTCTGAAATTT 6958
159 ----- 159
Db 6959 TTTTGTGTGAAGAAAGAGAGATGGCAAGAGTGTGAGTGTCTATAGGTCAAGAG 7018
160 ----- 161
Db 7019 TGGCGGCATCCCTCTGAGGTTCTCCACCCACCGCTTTTCACTCCACTCTGACAG 7078
161 yPheMetThrProAlaGluHisIleuGlnIleuGlnIleuSerLeuProHisAsnMetPh 181
161 CTTTATGACTCCGGCAGAACACAGACAGTGGAGAAACTGAGACCTTCCACACAGATGTT 7138
181 eTrpValProTrpValIleuPheAlaAsnLeuSerMetIleuValIleuGlnIleuValIleu 201
181 CTTTATGACTCCGGCAGAACACAGACAGTGGAGAAACTGAGACCTTCCACACAGATGTT 7138
181 eTrpValProTrpValIleuPheAlaAsnLeuSerMetIleuValIleuGlnIleuValIleu 201
201 eArgAspProIleuLeuGlnSerLeuAsn ----- 212
212 ----- 212
Db 7199 CCGGAGCCCTATCTGCTCCAGAGCTGCTGAACGTGAGCCCACTGTACAGACAGGCGTG 7258
212 ----- 212
Db 7259 CCGCAGAGTGGGAGGCGTGTGCTCCACAGAAACAGGTTTCTTACAAAGAGAACCTTT 7318
212 ----- 212

7319 GGGCCCTGAGGGTCTTCCGAGAGCTGAGTGGGGTTCAGAAATCTTTCCACAGCA 7378
212 ----- 212
Db 7379 TCCAGGCCGAGAGTGTCCCTTCTCAGAGGCCCTCCCTCTTCTCAAGTCTGTAGGT 7438
212 ----- 212
Db 7439 CCGTGTCCCTTTTGTATAGATGAGAGAGCTGAGACACAAAGAGTTTACTGACTTCCA 7498
212 ----- 212
Db 7499 TGGCCACAGCCAGGAATGACCATAGTACAGAGCCCTGGTACTGAGAAAGAGTGG 7558
212 ----- 212
Db 7559 GGGCGAGCCAGAGTGGGGGAGTGTGTTCAGAAACCCATCCCTCTCTGCCCCC 7618
213 --GluMetAsnThrIleuArgThrGlnCysGlnHisIleuTrpAlaTrpAspTrpIleSerI 232
213 AGGAGATGAAACACTTGGTGTACTGAGTGTGACACCTGTATGCTTACGACTGGATTAGTA 7678
232 IeProLeuValTrpThr ----- 237
237 TCCACTGGTGTATACACAGGTGAGAGCTAGGCTGTGAGGCTGCCCTTTTGGGAAACTG 7738
237 ----- 237
Db 7739 AGCTAGAAGACCAAGAAAGAGCTGGGGTGGAAAGGCTCAGTAGAGCTAACTGGC 7798
237 ----- 237
Db 7799 TCCCTGGAGATTGGGTCCACACTTGAAGTTGGGTGTGACTTGAAGTCCAGTTCT 7858
237 ----- 237
Db 7859 AAGATCCAGGCTCTGCTGGCCGAGTCCAGTAGAGCAATGTATATCCCATATTA 7918
237 ----- 237
Db 7919 AAGAGAGTTGGCGGGGTGACAGTGGCTCATGCTGAATCCAGACTTTGGGAAAGTGA 7978
237 ----- 237
Db 7979 GGCAGGTGATCATCTGAGTCAAGAGTTCGAGACCAAGCTGGCAACATGTGAAACC 8038
237 ----- 237
Db 8039 CATCTCTACTGAAATACAGAAATTAAGTGTGTGTGTGACAGCCTGTATATCCAGCTAC 8098
237 ----- 237
Db 8099 TTGGGAGGTGAGGCGAGAGAAATCGCTGAACCCGGAGGTGAGGTTGCAGTGAGCTGA 8158
237 ----- 237
Db 8159 GATCATGCCACTGACTCCAGCCTGGGCGACACAGCAAGACTGTGTCTCAAAACAAACAA 8218
237 ----- 237
Db 8219 CAACAAACAAACAAAGGGGTTAACAGAGCCCTTAAGTCAATTAAGTGTGCAAGTCAAA 8278
237 ----- 237
Db 8279 CAAGGCTTGTCTCTGTCTGAGACTCCAGGCCCTGAGCATCTGATTTCAGGGTTTC 8338
238 ----- 245
238 ----- 245
Db 8339 CCACATAGCCCTTGTGTACACATCCCTCTCTCTCCAGGTGTGACTGTGGCGGTGT 8398
245 ySerIleuPheLeuIleuTrpCysLeuValGlnIleuAsnProAlaIleuValTrp 265
245 ----- 265

Db 8399 ACAGCTCTCTGCTGACTGTGTAGTTGGGCGGCACTTTCTGAAACCCAGCAAGCCCTTACC 8458
Oy 265 roglYhISgluLeuaspLeuValProValPheThrPheLeuGlnPhePhePheTyY 285
Db 8459 CTGGCAGAGAGGTGACCTCGTTGTGCCGCTTACAGCTCTGCAAGTCTCTCTCTATG 8518
Oy 285 aIGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeuGlyTrpArgHISglYglnAATg 305
Db 8519 TTGGCTGGCTGAAGGTGGGCTCTCCAGGCCCTCTGGGCTGGAGGCAATGGCCAGAGG 8578
Oy 305 lYhISglYglnLeuLeuGluTrpArgMetGlnCysGlnGluArgLysValSerArgY 325
Db 8579 GTCATGGCCAGAGGTGCTGCTGACGAGATGATGATGAGAAAGAGGTCTCAGGG 8638
Oy 325 aIGuSerSerGlnAlaTrpTrpArgThrProValIleProAlaThrArgLysAlaGlnA 345
Db 8639 TGAAGAGGAGCCAGGCGTGTGGCGACACCTGTATCCAGCTACTCGGAGGCTGAGG 8698
Oy 345 lAGlyGluSerLeuGluProGlyArgArgLeuTrp----- 357
Db 8699 CAGGAGAAATCGCTTGAAACCCGGAGGCGGAGGTTGTGAGTTGAGATCGTCCACTGC 8758
Oy 357 ----- 357
Db 8759 ACTCCAGCTGGGCAAAAGATGAATCTATCTCAAAAAACAACAACAACAACAACA 8818
Oy 357 ----- 357
Db 8819 AGCCCTAAGGTTAGAAAGCCCTGCCCTTAGAAGAGATGCGCACACTCTCTAT 8878
Oy 357 ----- 357
Db 8879 TCAAGATGCTGTGGGCTGTCTTGTCTCCTCAGCTCAAGTGGCTTCCAGATTCCT 8938
Oy 357 ----- 357
Db 8939 CCCACACAGCAATACCTCGAACAAGATGTTGTAATCAACAGTTTCTCCACTCT 8998
Oy 357 ----- 357
Db 8999 TTATCTTCTCTCTCTGTGTGCCACACACTCTCTCTCTCTCTCTCTCTCTAT 9058
Oy 357 ----- 357
Db 9059 TTTGGTAATGGGGGTAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9118
Oy 357 ----- 357
Db 9119 CACACACACACACACACACATACACACACACACACAGCATTTCTATTCCTTAAT 9178
Oy 357 ----- 357
Db 9179 CCCCCTGCCCCCACTTATCTTGTGTTCTGAGATCAAAACAAATCACACTTTATGC 9238
Oy 357 ----- 357
Db 9239 TTGAATTTCCGGGGTCCCACTGAGTCCCAAGATGTCCCTGGAGCCCTAAGGAGA 9298
Oy 357 ----- 357
Db 9299 CGGGTGTACACTCTGCGGGCTTTGTTAGGGCATTTTATAGAGTTGATCCAGAAATCTG 9358
Oy 357 ----- 357
Db 9359 CCCACCTAGACTGCCCTTATGTTACGCCCACTTCACTATATATCTCTGTTCATGAATG 9418
Oy 357 ----- 357
Db 9419 AATTAATTTATGCAACTCAGAGTAAGATGAGATGAGATAAAGCAGTACTCAGCC 9478
Oy 357 ----- 357
Db 9479 GAGTATACACTCAGGAGACAGCTGTGGTGTTCAGGGAAGACGTGCTCAGAAAGTTAG 9538

Oy 357 ----- 357
Db 9539 AGGGCTGTGTCCAGAAAGTGTGTGGTGGCCCAAAAGTGTGGGGCTGGAGCCCTAACT 9598
Oy 357 ----- 357
Db 9599 CTGGCTTTGAAGACAGTGGTGCAGGAGGAAGGCGTATGGGTGTGAATATAGCAGAG 9658
Oy 357 ----- 357
Db 9659 CTGAGTTTAAAGGGGAAGCTGGCTTGAAGAGTTCTGCTGAGGGTTTACAGACCCTC 9718
Oy 358 ----- 373
Db 9719 ACCTGCCCCAAGGTGGAGACAGCTCTATCAACCCCTTTGGAGAGATGATGATATT 9778
Oy 373 euArgProThrGlyLeuSerThrGlyLeuCys----- 383
Db 9779 TGAGACCAACTGGATGTGACAGGAATTTGCAGATGAGGAGAGAGAACCAAA 9838
Oy 383 ----- 383
Db 9839 CCATGGACCTTCCCAAAAGTGAACCAAGAGAGACCCCACTGTTCTGAGGAGCCT 9898
Oy 383 ----- 383
Db 9899 CACAGTGAATGATCAACCTTCCCTCTCTCTCTCTGACGACGATTCATCAGAGAG 9958
Oy 383 ----- 383
Db 9959 TTCTCAGCTCAATCTTTGAGGCTGACGAGGAGCACCACATCTCCATTTCACAGGAGG 10018
Oy 383 ----- 383
Db 10019 AAAGTGAAGTCCAGAGAGAGAGAGATTCCTCAAGTCAAGCAGCATACAAAGTCT 10078
Oy 384 ----- 388
Db 10079 GCCTGGAGATGATCTTGTGGGACTTCTCTGCTCCCTGTGACAGAGTCTCCCTGTGG 10138
Oy 389 LeuTrpMetArgCysThrArgThrCysLeuGlyTrpSerArgThrCysThrArgYlIeser 408
Db 10139 CTGTGATGATGATGACACAGAGCTCTGATGAGGCGGACATGATCTGGAATATAGC 10198
Oy 409 ProSerHisSerProProThrGlnLeuLeuProProSerSerValGluProProLeuTrp 428
Db 10199 CCGAGCCACAGCCCTTACACAGCTGCTTCCGCGCCAGTTCCTGCGAGCTCTTATGG 10258
Oy 429 AlaProSerThrSer 434
Db 10259 GCTCACACTTCACATCA 10276
RESULT 6
AC003025 112309 bp DNA linear PRI 23-JUL-1998
LOCUS Human Chromosome 11p12.2 PAC clone pdv466a1, complete sequence.
DEFINITION AC003025
ACCESSION AC003025.1 GI:337308
VERSION
KEYWORDS
SOURCE HTG.
ORGANISM Homo sapiens.
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basil,M.,
Buetner,J., Bumeister,R., Card,P., desailloat,F., Dunn,J.,
English,C., Elnridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,T., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
HTGS Submission
TITLE
JOURNAL
Unpublished

REFERENCE 2 (bases 1 to 112309)
AUTHORS Evans,G.A., Athanasiou,M., Bradbury,P., Brignac,S., Bumeister,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Maryanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
REFERENCE 3 (bases 1 to 112309)
AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Bumeister,J., Bumeister,R., Card,P., Desaliboat,F., Dunn,J., English,C., Ehrldge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA
COMMENT On Jul 23, 1998 this sequence version replaced g1:2554967.
IMPORTAT: This submission contains the entire insert of clone PDJ4661. PDJ4661 comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected Phred/Phrap calculated errors/10kb is 0.18. In addition, this sequence has been finished such that 99.9% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.
Further information regarding the map of this region or annotation of PDJ4661 can be found at <http://gestec.smed.edu/chromosol.htm>
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST AHNAK. This region spans over 1.5 Mbp.
MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA)
MAPPED CLONE OVERLAP: HTGS submitted PAC clones pdj519c13 and pdj756b9.
FEATURES
source
1. .112309
/organism="Homo sapiens"
/db_xref="taxon:9606"
complement(4341..4656)
repeat_region
/rpt_family="Alu"
complement(9880..10120)
repeat_region
/rpt_family="Alu"
complement(10163..10282)
repeat_region
/rpt_family="Alu"
complement(11232..12157)
repeat_region
/rpt_family="Alu"
complement(13351..13560)
repeat_region
/rpt_family="L1"
complement(15092..15146)
repeat_region
/rpt_family="MIR"
complement(15826..15932)
repeat_region
/rpt_family="MIR"
complement(17011..17146)
repeat_region
/rpt_family="MIR"
complement(18018..18313)
repeat_region
/rpt_family="Alu"
complement(18493..18741)
repeat_region
/rpt_family="Alu"
complement(18820..18954)
repeat_region
/rpt_family="Alu"
complement(21374..21661)
repeat_region
/rpt_family="Alu"
complement(22809..23212)
repeat_region
/rpt_family="Alu"
complement(23499..23752)
repeat_region

repeat_region
/rpt_family="Alu"
complement(24118..24472)
repeat_region
/rpt_family="Alu"
complement(25005..25113)
repeat_region
/rpt_family="MIR"
complement(31160..31442)
repeat_region
/rpt_family="Alu"
complement(32651..32947)
repeat_region
/rpt_family="Alu"
complement(35970..36025)
repeat_region
/rpt_family="MIR"
complement(38958..39145)
repeat_region
/rpt_family="Alu"
complement(44878..45480)
repeat_region
/rpt_family="Alu"
complement(49769..50071)
repeat_region
/rpt_family="Alu"
complement(50172..50310)
repeat_region
/rpt_family="Alu"
complement(50902..50982)
repeat_region
/rpt_family="MIR"
complement(51038..51659)
repeat_region
/rpt_family="Alu"
complement(54494..54769)
repeat_region
/rpt_family="Alu"
complement(55529..55623)
repeat_region
/rpt_family="THF1"
complement(55967)
repeat_region
/rpt_family="Alu"
complement(56049..56257)
repeat_region
/rpt_family="Alu"
complement(56721..57100)
repeat_region
/rpt_family="THF1"
complement(58402..58680)
repeat_region
/rpt_family="Alu"
complement(58748..59345)
repeat_region
/rpt_family="Alu"
complement(59678..59904)
repeat_region
/rpt_family="Alu"
complement(61149..61424)
repeat_region
/rpt_family="Alu"
complement(61930..62242)
repeat_region
/rpt_family="Alu"
complement(62276..62710)
repeat_region
/rpt_family="MIR1"
complement(62915..63118)
repeat_region
/rpt_family="Alu"
complement(63349..63455)
repeat_region
/rpt_family="Alu"
complement(64339..64614)
repeat_region
/rpt_family="Alu"
complement(64707..64984)
repeat_region
/rpt_family="Alu"
complement(65596..65811)
repeat_region
/rpt_family="Alu"
complement(66500..66753)
repeat_region
/rpt_family="Alu"
complement(67311..67542)
repeat_region
/rpt_family="Alu"
complement(67835..68119)
repeat_region
/rpt_family="Alu"
complement(68860..69150)
repeat_region
/rpt_family="Alu"
complement(69164..69272)
repeat_region
/rpt_family="Alu"
complement(69354..70245)
repeat_region
/rpt_family="Alu"
complement(70655..70939)
repeat_region
/rpt_family="Alu"
complement(70953..71075)
repeat_region
/rpt_family="L1"
complement(71119..71395)
repeat_region
/rpt_family="Alu"

```

repeat_region      71404..71530
/rpl_family="LTR5"
repeat_region      71536..71637
/rpl_family="LTR5"
repeat_region      72720..72948
/rpl_family="L1"
repeat_region      complement(73488..73633)
/rpl_family="Alu"
repeat_region      complement(73925..74200)
/rpl_family="Alu"
repeat_region      75721..75955
/rpl_family="Alu"
repeat_region      77185..77449
/rpl_family="Alu"
repeat_region      78093..78322
/rpl_family="L1"
repeat_region      78227..78294
/rpl_family="L1"
repeat_region      complement(80458..80707)
/rpl_family="Alu"
repeat_region      complement(81544..81974)
/rpl_family="Alu"
repeat_region      82055..82305
/rpl_family="Alu"
repeat_region      82514..83096
/rpl_family="Alu"
repeat_region      complement(83685..83940)
/rpl_family="Alu"
repeat_region      85336..85850
/rpl_family="Alu"
repeat_region      complement(86429..86483)
/rpl_family="MIR"
repeat_region      86902..87183
/rpl_family="Alu"
repeat_region      87646..87792
/rpl_family="Alu"

```

Alignment Scores:

Pred. No.:	8,016-71	Length:	112309
Score:	1302.00	Matches:	383
Percent Similarity:	23.17%	Conservative:	0
Best Local Similarity:	23.17%	Mismatches:	1
Query Match:	55.40%	Indels:	1269
DB:	9	Gaps:	6

```

US-09-622-964-5 (1-435) x AC003025 (1-112309)

QY 51 ATGLeuAlaLeuThrGluGlnGlnLeuMetPheGluLysLeuThrLeuTyrCysasp 70
   |||||
DB 84312 AGCGTGGCCCTCAGCGAAGAACACACAGCTGATGTTGAGAAACTGCTGTATTGCGAC 84371

QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
   |||||
DB 84372 ACCTACATCCAGCTCATCCCATTTCTTCGTGCTGGTGAAGTCTCCCTCTGCTGTT 84431

QY 82 ----- 82
DB 84432 CCGGGTCCCTGTGGCCGCCAGGCTCCACACAGGCCAGGGAGGATCAGAGGAGCTGGG 84491

QY 82 ----- 82
DB 84492 GCAAGGGCTGGGAGGCGGGGAGAACGCCAGCGGAGTGGCCCTCTGTAGGG 84551

QY 82 ----- 82
DB 84552 AAAGTGGAGCTGACGACGAGAAACTGAATTAGAGTTAGGTAAAGAGCTCTGCCGT 84611

QY 82 ----- 82
DB 84612 TAGCATGAAAAACCCCATTTTCTGAGGGAAGCGCTACATCATGTCCTGGAGCCCTG 84671

QY 82 ----- 82

```

```

DB 84672 CGCGGAGGAGGAGGGGCTGTGGCGATTTCTGGGACACGAGGGGGAACCCCGGGTGACA 84731
QY 82 ----- 82
DB 84732 GAACCTTGGGGCTCTCGCGCCCTCCATGAGGCTGTGCTCTGCTCCGAGCGC 84791
QY 82 ----- 82
DB 84792 CTTCCAGAGGGCTGGGGGCTAGGCCCGCTGCAGAGAAAGCTGAGAGAGCCAGGCAT 84851
QY 82 ----- 82
DB 84852 CGCGGGCGCTGGGCCCTGGGCTGTGGCCGACCTGGCCCTGGCCCTGGCCCGCCG 84911
QY 83 ----- 83
DB 84912 CCCTCTGCCAGGCTTCTAGTAGAGCTGTGCTGAGACCCGCTGGTGAACAGTACAG 84971
QY 99 AsnLeuProTrrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAsp 118
   |||||
DB 84972 AACCTGCGGCGCCGACCGCTCATGAGCTGTGTGGGCTTCTGTGAAGGCAAGGAC 85031
QY 119 GluGlnGlyArgLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
   |||||
DB 85032 GAGCAAGGCGGCTGCTGGCGGACGCTCATCCGCTACGCCAACCTGGGCAACGTGCTC 85091
QY 139 IleLeuArgSerValSerThrAlaValTyrIleArgPheProSerAlaGlnIleLeuVal 158
   |||||
DB 85092 ATCTGCGCCACGCTGTAGACCGCACTTACAAAGGCTTCCCGAGCGCCAGCAGCTGGTG 85151
QY 159 Gln----- 159
DB 85152 CAACGAGTGGGCGGACCGGAGCAACGAGGACCGGCGAGCCAGAGGCGCGAGAT 85211
QY 159 ----- 159
DB 85212 GGGCGGCGAGATGAAATGGGTGGAGCCAAAGTCCCGGACTGGGGGATTGGGT 85271
QY 159 ----- 159
DB 85272 GAGGCGAGAGTGGGGTGTGATCAAGATTGGGGGTCCAAATGGCGGAGAGTGGG 85331
QY 159 ----- 159
DB 85332 TGTCTGAAGTGGGCGAGGCCAGGACCCACCTCCGAGAGTGAAGTGTGAGCGAGG 85391
QY 159 ----- 159
DB 85392 CTAAAGACCTTTAGAGGATATGAAAGAGGCTGACGGCTTGGAACTGTGAGTACT 85451
QY 159 ----- 159
DB 85452 AAGGTCTACTTCCCTGCTGCCCCCTTGCCCCCTTGATCTCCGTTTCCACTGTGAGGTATG 85511
QY 159 ----- 159
DB 85512 GGACATTGGTCTGTGACACCCCTCAGCCTGGCCTGACTGTGTTAATAGACAG 85571
QY 159 ----- 159
DB 85572 ACCCAGGCTAGGCGTGTGTGCTGTGCTGTATCCCACTGCTTTAGAGGCAAGGTGG 85631
QY 159 ----- 159
DB 85632 GAAGATCGCTTAGGCCACGCTGTTGAGAGCCCTGAGCAACATAGCAGAGACCCCATC 85691
QY 159 ----- 159
DB 85692 TCTACAAAAACATTAAAAATTAGCAGGAGATGCTGCTGTGCTGTAGTCTGAGGCTGA 85751
QY 159 ----- 159
DB 85752 GTATCGGAGGCTGAGGCGAGAGATCACTTGAGCCACAGTTCAGAGGCTGCAGTGCGG 85811

```

Oy	159	-----		-----	159
Db	85812	TAAAGTCGACCGCTGCACCTCCAACTCGGTGACAGAGCCAGACCCTTTCTCTGGAAATA			858711
Oy	159	-----		-----	159
Db	85872	AATTAATACCTGCCCATGCTCAGCCCCAGAACACACTAGTAGTGCTCAGAAATT			859311
Oy	159	-----		-----	159
Db	85932	TTTTGTTGTAAGAAAAGAGATGGCAAAGAGTGTGAGTTCTATAGTTCAGACAG			859911
Oy	160	-----	AlaGI	161	
Db	85992	TGCCGGCCANCCCTTCGACGGTTCTCCACCACCGCCTTCTTCACTCCACTCTGCAG			860511
Oy	161	yPheMetThrProAlaGluHisLysGlnLeuGlyIuLysSerLeuProHisAsnMetPh			181
Db	86052	CTTTATATACGTCGGGCAACAACAGACAGTGGGAAGAACAGAGCTCACACACATGTT			861111
Oy	181	eTrpValrProTrpValrTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyGlyArgIl			201
Db	86112	CTGGGTCCTCGGGGTGGTGGTTGCCAACCCTGCAATGAAGGCGCTGGAGTGCAGAT			861711
Oy	201	eArgAspProIleuLeuGlnSerLeuLeuAsn-----			212
Db	86172	CGGGGACCCATCTGCTGCTCAGAGCCGCTGAACGTGAGCCCACGTACAGACAGGGCTG			862311
Oy	212	-----			212
Db	86232	CCGCAAGTGGGAAGGGTTGGTCCACAGAAACAAGTTTCTACAAAGAAGACCTT			862911
Oy	212	-----			212
Db	86292	GGGCCCTGAGGCTTTCGAGAGCGGAGTGGGGTTGCAGATCTTTCACAACAGCAA			863511
Oy	212	-----			212
Db	86352	TCCACAGCCGAGTGGTCCCTTCTCAGAGGCCCTCCCTTCTCCAAGTCTGTAGGT			864111
Oy	212	-----			212
Db	86412	CCTGGTTCCCTTTGATAGATGAGGAAGCTGAGACACAAGAAGAGTTTAGAGCTTCCCA			864711
Oy	212	-----			212
Db	86472	TGGCCACACAGCCAGGAATGACCATAGTACCAAGGCCCTGTACTGGAGAGAGTGG			865311
Oy	212	-----			212
Db	86532	GGGGAGACCCAGGGTGGGGGAGGTGGTGTTCAGAAACCCCATCCCTCTTGCCCCC			865911
Oy	213	--gluMetAsnThrLeuArgTrgHincysGlnHisLeuTrpAlaTrpAspTrpIleSerI			232
Db	86592	AGGAGAGGAACACTTCTCGTACTCAGTGTGAGACACCGTATGTGCTAGCACTGATTAAGTA			866511
Oy	232	IeProLeuValrTyrrThr-----			237
Db	86652	TCCCACTGCTGTATACACAGAGTGAAGACTAGGCTGTGAGGCTGCCCTTTTGGGAACCTG			867111
Oy	237	-----			237
Db	86712	AGGCTAGAGAGCAACAGGAAGACAGCTGGGGTGGGAAGGGCTCACCTTAGAGGCTTAAGTGGC			867711
Oy	237	-----			237
Db	86772	TCCCTGGGATTTGGGTCCACACTTTGAAGTTGGGTGTGACTTTGAAGTGCACAAGTTCT			868311
Oy	237	-----			237
Db	86832	AAGAGTCAGGCTCCTGCTGGCCCACTGCAAGTAGAGCAATGTGATTATCCCAATATTA			868911

[illegible]


```

repeat_region      /rpt_family="MIR"
4417. .4995
/rpt_family="Alu"
complement(4432. .4936)
repeat_region      /rpt_family="SVA"
complement(5503. .6111)
repeat_region      /rpt_family="Alu"
5569. .6095
/rpt_family="SVA"
complement(7239. .7542)
repeat_region      /rpt_family="Alu"
7306. .7497
/rpt_family="SVA"
7752. .7828
/rpt_family="MIR"
9099. .9403
/rpt_family="Alu"
complement(9526. .9897)
repeat_region      /rpt_family="MER1"
9635. .9752
/rpt_family="MER1B"
11924. .12197
/rpt_family="Alu"
complement(11955. .12151)
repeat_region      /rpt_family="SVA"
complement(12767. .13466)
repeat_region      /rpt_family="Alu"
16757. .17039
/rpt_family="Alu"
complement(16775. .17010)
repeat_region      /rpt_family="SVA"
complement(18220. .18507)
repeat_region      /rpt_family="Alu"
18229. .18485
/rpt_family="SVA"
20101. .23446
/gene="HDMFRITR"
/note="Human ferritin heavy chain gene"
join(20101. .20421,22218. .22364,22621. .22746,22842. .23446)
/gene="HDMFRITR"
/product="ferritin heavy chain"
complement(21466. .21731)
repeat_region      /rpt_family="Alu"
21506. .21714
/rpt_family="SVA"
complement(23272. .37375)
/gene="VMD2"
/note="Best macular dystrophy gene"
complement(join(23272. .23629,24840. .25478,27690. .27841,
28155. .28235,29427. .29579,30274. .30351,30740. .30894,
31787. .32020,32537. .32631,35777. .35964,37308. .37375))
/gene="VMD2"
/product="Best macular dystrophy protein"
complement(23735. .23934)
repeat_region      /rpt_family="Alu"
23789. .23956
/rpt_family="SVA"
26429. .26707
/rpt_family="Alu"
complement(26444. .26696)
repeat_region      /rpt_family="SVA"
27103. .27349
/rpt_family="Alu"
complement(27106. .27323)
repeat_region      /rpt_family="SVA"
complement(29161. .29307)
repeat_region      /rpt_family="Alu"
complement(29762. .30043)
repeat_region      /rpt_family="Alu"
29786. .30017
/rpt_family="SVA"
30462. .30516
/rpt_family="MIR"

```

```

repeat_region      complement(31095. .31409)
/rpt_family="Alu"
33005. .33260
/rpt_family="Alu"
33247. .33597
/rpt_family="LMC2"
complement(33849. .34431)
repeat_region      /rpt_family="Alu"
34076. .34411
/rpt_family="SVA"
complement(34640. .34890)
repeat_region      /rpt_family="Alu"
34972. .35401
/rpt_family="Alu"
complement(34988. .35392)
repeat_region      /rpt_family="SVA"
36238. .36487
/rpt_family="Alu"
complement(36253. .36457)
/rpt_family="SVA"
complement(38661. .38855)
/rpt_family="L1"
complement(40997. .41224)
repeat_region      /rpt_family="Alu"
41056. .41213
/rpt_family="SVA"
42745. .43020
/rpt_family="Alu"
complement(42763. .42966)
repeat_region      /rpt_family="SVA"
43312. .43457
/rpt_family="Alu"
complement(43963. .44217)
/rpt_family="L1"
complement(45296. .45533)
/rpt_family="SVA"
complement(45300. .45401)
/rpt_family="LFR5"
complement(45407. .45533)
/rpt_family="LFR5"
complement(45542. .45818)
repeat_region      /rpt_family="Alu"
45598. .45770
/rpt_family="SVA"
complement(45862. .46004)
/rpt_family="L1"
complement(45998. .46282)
/rpt_family="Alu"
46022. .46260
/rpt_family="SVA"
complement(46692. .47583)
repeat_region      /rpt_family="Alu"
46885. .47215
/rpt_family="SVA"
47363. .47571
/rpt_family="SVA"
complement(47665. .47773)
/rpt_family="Alu"
complement(47792. .48077)
/rpt_family="Alu"
complement(48110. .48388)
/rpt_family="L1"
complement(48827. .49102)
repeat_region      /rpt_family="Alu"
48883. .49110
/rpt_family="SVA"
complement(49395. .49626)
/rpt_family="Alu"
50184. .50437
/rpt_family="Alu"

```

Alignment Scores:
Pred. No.:

3.86e-70

Length:

142092

Score: 1293.00 Matches: 383
Percent Similarity: 23.20% Conservative: 0
Best Local Similarity: 23.20% Mismatches: 1
Query Match: 55.02% Indels: 1267
DB: 9 Gaps: 6
US-09-622-964-5 (1-435) x AF139813 (1-142092)
OY 51 ArgLeuAlaLeuThrGluGluGlnLeuMetPheGluLysLeuThrLeuTyrCysAsp 70
DB 32633 AGGCGGCGCCCTCACGAGAAACACAGCTGATGTTGAGAAACTGCTGATTCGAGC 32574
OY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu----- 82
DB 32573 AGCTACATCCAGCTCATCCCATTTCTCTCTGTCGTGAGTGTCCCTCTGCGCTGT 32514
OY 82 ----- 82
DB 32513 CCGGGTCCTGTGGCCGCCAGGCTCCAGACAGAGGAGAGATCAGAGAGCTGCG 32454
OY 82 ----- 82
DB 32453 GCAGGGGCTGGGAGGGGGCGGGGAGCCAGCGAGGTCGGCCTCTCTGTAAGG 32394
OY 82 ----- 82
DB 32393 AAAGTGGCGACTGCACGACAGAAACTGAATTAGACTTGAAGAGCTCTGCGCT 32334
OY 82 ----- 82
DB 32333 TAGCATGAAACCCCATTTTCTGAGGGAAGCGCTGACATCATGTCCTCGAGCCCTG 32274
OY 82 ----- 82
DB 32273 CCGGGAGGGAGGGGGTCTGGCGATTCTGGACAGAGGGGAGCCCGGGGTGACA 32214
OY 82 ----- 82
DB 32213 GAACTTGGGGTCTCGGGCCCTCCATGAGAGCTCTGCTGCTCTCCGAGCGC 32154
OY 82 ----- 82
DB 32153 CTTCAGAGAGGCTGGGGCTAGGCCGCTCGCAGACAGAAAGCTGAGAGCGAGCAT 32094
OY 82 ----- 82
DB 32093 CCGCGGGCGCTGGCCCTGCGCTCTGCGCGAGCTGGCCCTCGCCCTCGCCCGCCG 32034
OY 83 -----GlyPheTyrValThrLeuValValThrArgTyrTrpAsnGlnTyrGlu 98
DB 32033 CCTCTCGCCCAAGGCTTCTACGTGACGCTGTGTCGACCGCTGTGTGAACAGTAGAG 31974
OY 99 AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
DB 31973 AACCTCGCGGCGCCGACCCCTCATAGAGCTGTGTGCGGGCTCTCTCGAAGGCAAGAC 31914
OY 119 GluGlnGlyArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
DB 31913 GAGCAGAGCGGGCTGTGCTGCGCGCACGCTATCCGCTACGCCCAACTGGGCAACGTCCTC 31854
OY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal 158
DB 31853 ATCTCTGGCGAGCTGACACCGCAGTCTACAAGGCTTCCCGACGGCCAGACACTGTGTG 31794
OY 159 Gln----- 159
DB 31793 CAGAGAGTGGGGGAGACCGGAGACAGGGAGGACCGGGCAGAGAGCGCGAGAT 31734
OY 159 ----- 159
DB 31733 GGGCGCGCAGAGATGAGAGATGGGTGAGCAAAAGTCCCGGAGCTGGGGGATTGGGT 31674
OY 159 ----- 159

DB 31673 GGAGCCAGAGTGGGGTGTGTCCAGATTTGGGGGTCCAATTGGCGGGACAGAGTCGGG 31614
OY 159 ----- 159
DB 31613 TGTCTGAGAGTGGGGGAGGCGAGAGCCACCTCCGAGAGTGAAGTCTGAGGCAAGG 31554
OY 159 ----- 159
DB 31553 CTAGAGACCTTGAAGGATTAATGAAAGAGGTGAAGGCTTGGAACTGTGAGTACT 31494
OY 159 ----- 159
DB 31493 AGGCTACTTCCCTCTGCTGCCCTTGGCCCTTGATCTCCGGTTCACCTGTGAGGTATG 31434
OY 159 ----- 159
DB 31433 GGACATTGTCTCTGACACCCCTCAGCCTGACCTGCTGCTGTGTTAATTAAGACAG 31374
OY 159 ----- 159
DB 31373 ACCAGGCTAGGGGTGTGCTCTCGCTGTAAATCCAGTGTAGAGGCAAAAGGTGG 31314
OY 159 ----- 159
DB 31313 GAAGATCGCTTGAGCCAGCTGTTGAGACGCCCTGAGACATAGCGAGACCCCATC 31254
OY 159 ----- 159
DB 31253 TCTACAAAACATTAAATTAACAGGCGATGTGTGCTGTGCTGATGCTGAGCTGA 31194
OY 159 ----- 159
DB 31193 GTATCGGAGGCTGAGGCGAGAGGATCATTTGAGCCAGAGTTCAGGCTGAGTGGC 31134
OY 159 ----- 159
DB 31133 TAAGATCGACACGCTGCTCAACCTCGGTGACAGAGCAGACCTTTCTGTGAATA 31074
OY 159 ----- 159
DB 31073 AATTAATACCTTGCCTCAGCTCAGCCAGACAGCACTAGTAGGTGCTCAGAAATTT 31014
OY 159 ----- 159
DB 31013 TTTTGTGTGAAGAAGAGATGCCAAGAGTGTGAGTCTATAGTCAAGCAG 30954
OY 160 -----AlaGln 161
DB 30953 TGCGGCCATCCCTTCTGAGGTTCTCCACCCAGCCGCTTCTTCACTCCACTGTGAGG 30894
OY 161 YPheMetThrProAlaGlnHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetLph 181
DB 30893 CTTTATGACTCCGACACACACAGAGCTGTGAATACTGAGCTTACACACACATGTT 30834
OY 181 eTyrValProTyrValTrpPheAlaAsnLeuSerMetLysAlaTyrLeuGlyArgIle 201
DB 30833 CTGGGTCCCTCGGGTGTGTTGGTTCACACCTGTCAATGAAGCGTGTGGAGGTGCAAT 30774
OY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn----- 212
DB 30773 CCGGACCTATCTGTCTCCAGAGCTGTGTAAGCTGAGCCCACTGTACAGACAGGCTG 30714
OY 212 ----- 212
DB 30713 CCGCAGAGTGGGAAGGTTGTGTCACAGAAACAGGTTCTCTACAAGAGAGCTT 30654
OY 212 ----- 212
DB 30653 GGGCTCTGAGGCTTTCGAGAGCCGAGGTGGGTTGCAGAACTTTTCCACAGCA 30594
OY 212 ----- 212

Db	30593	TCACACGCCCGAAGGTGGTCCCTTCTCAGAGGCCCTCCCTCTTCTCCAGTCTGTGAGGT	30534
OY	212	-----	212
Db	30533	CCTGGTTCCTTTTGGATGATGAGGAACCTGAGACACAAAGAGTTTAGTAGCTTCCCA	30474
OY	212	-----	212
Db	30473	TGGCCACACAGCCGAGATGACCATAGTACACAGGCCCTGGTACTGGAGAGAAGGTGG	30414
OY	212	-----	212
Db	30413	GGCGAGCCCGAAGGTGGGGGAGGTGGTGTTCAGAACCCATCCCTCTCTTGCCCCC	30354
OY	213	-GluMetAsnThrLeuArGhrGlnCysGlyHisLeuTyrAlaTyrAspIleSerI	232
Db	30353	AGGAGATGAACACCTTCGCTACTCAGTGTGACACACTGTATGCTCAGACGTGATTA	30294
OY	232	IleProLeuValTyrThr-----	237
Db	30293	TCCCACTGTGTATACACAGGTGAGACTAGCTGTGTGAGGCTGCCCTTTGGGAAATG	30234
OY	237	-----	237
Db	30233	AGGCTAGAACGACCAAGAAAGCAAGCTGGGTGGGAAGGGCTCACCTAGAGCTAAGTGGC	30174
OY	237	-----	237
Db	30173	TCCCTGGAGATTGGGTCCACACTTTGAAGTTGGGTCTGGACTTTGAAGTCCCAATTCT	30114
OY	237	-----	237
Db	30113	AAGAGTCCAGGCTCTGCTGGCCAGTCCAGTAGAGCAATGTGATTATCCCATATTA	30054
OY	237	-----	237
Db	30053	AAGAGAGTTGGCGGGGACAGTGGCTCATGCTGTATCCACGACTTTGGGAAGCTGA	29994
OY	237	-----	237
Db	29993	GGCAGGTGATCATCTGAGGTCAAGATTGAGACCAAGCCTGGCCACATGTGTAAACC	29934
OY	237	-----	237
Db	29933	CATCTCTACTGAATAATACAGAAATTAGCTGTGTGTGTGCACGCTGTAAATCCAGCTAC	29874
OY	237	-----	237
Db	29873	TTCGGAGGCTGAGGCGACGAGAAATCGCTTGAACCCGGAGAGGTGAGATTGCAGTAGCTGA	29814
OY	237	-----	237
Db	29813	GATCATGCGACTGCACTCCAGCCTGGCGCACACAGCAAGACTGTGTCTCAACAAACAAA	29754
OY	237	-----	237
Db	29753	CAAAACAACAAACAACAAGGGGTTACAGAGCCCTAAGTACATTAAGTGTCAAGTC	29694
OY	237	-----	237
Db	29693	AGAACAGGCCCTTGGTCTCTCTCTCAGACTCCACGCCCTGGAGGATCTGTATTCAGG	29634
OY	238	-----GlnValIval	240
Db	29633	GTTCCACCTAGCCCTTTCCTACACACATCTCTCTCTCTCCCTCTCTCTCCACAGGTGGTG	29574
OY	241	ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro	260
Db	29573	ACTGTGGGGGTGTACAGCTTCTCTCTGACTTGTAGTTGGGGGCGAGTTTCGAACCCA	29514
OY	261	AlaLysAlaTyrProGlyHisGluLeuAspLeuValIvalProValPheThrPheLeuGln	280
Db	29513	GCCAAAGGCTACCTCGGCCATGAGGTGGACCTGTGTGGCCGTCTTCACTGCTTCGACG	29454

OY	281	PheherberetryvalglYTrpleuylsvalglyleUserArghAlaleuenglYTrparg	300
Db	29453	TTCCTTCTATGTTGGGTGCCTGAAGTGggccCTCCAGGGCCCCCGCTGGAGG	2939
OY	301	HISglYlnArGlYHISglYglNlInleuenglUThrArgMecIncysgnIunArg	320
Db	29393	CATGCCAGAGGGGCTCATGGCCACGCACTGCTGTAGACGAGATGCAgtGTcAGGAAGG	2933
OY	321	LysValSerArgValglUserSercInalatrYrPargThrProvalIlleProalatrHr	340
Db	29333	AAGGTCTACAGGGTAGAAAGCAGCCAGGCGTGTGGCGCACACTTAATCCAGCTACT	2927
OY	341	ArgLlnAgluaIaglyglUserleuglUprogIyArGarArgLeutrp-----	357
Db	29273	CGGAGGCTGAGCGAGGAGTAATCGCTTGAACCCGGAGGCGGAGGTGGTGAGTTGAG	2921
OY	357	-----	357
Db	29213	ATGTCGCACCTGCACCTCCAGCCTGGGCAAAAAGATGAACCTCTATCTCAAAAAACAAC	2915
OY	357	-----	357
Db	29153	AAACAACAAAACAAAGCCCTAAGTTTCAGAACCCCTGCCCTTTAGAAGGACAGCGCAC	2909
OY	357	-----	357
Db	29093	CACCTCTCTATTAAAGATGCTGTGGGCTGTCTGTCTCTCACCTCAAGTGCTGT	2903
OY	357	-----	357
Db	29033	CCAGGATTTCCCTCCACCACACAGCCAATACTCCGACACAGATGTTCTGAATCACACAGTTT	28974
OY	357	-----	357
Db	28973	TCTCTCACCTTTATCTTCTCTCTCTCTGTGTCGCCACCACTCTCTCTCCCTCA	28914
OY	357	-----	357
Db	28913	CCCTTCCTTATTTTTTGGTAATGGGGGTGAAGTCTGTCTGTCTGCTCCCTGTCACTG	28854
OY	357	-----	357
Db	28853	TGACATTCCTATTCCT	28794
OY	357	-----	357
Db	28793	AAATTCCTCCCTGCCCCCCCAGTATATCTTGGTTTCTGCAATCAAACAATACACATTT	28734
OY	357	-----	357
Db	28733	TATGTTGAATTTCTCCAGGTCGCCCACTGGGCTGCAGATGTCCTTGACCCCTAAG	28674
OY	357	-----	357
Db	28673	GCAAGCGCGTGTACCTCTTCGGGCTTGTGTAGGCAATTTAGAGTTGCTATCCAGGA	28614
OY	357	-----	357
Db	28613	ATCTGCCCACTAGACTGCCCTTTAGTTCAAGCCAGCTTCAGATATATCTCTGTGCAT	28554
OY	357	-----	357
Db	28553	GNATGAATAAATTTATGAACCTCCAGGTAAAGATACATGAGGTAGATAAAGCAGTGACT	28494
OY	357	-----	357
Db	28493	CAGCCGAGTATACACTCAGGACACAGCTGTGGTGTTCAGGGAAGACGTGGCTCAGAGA	28434
OY	357	-----	357
Db	28433	GTTAGAGGGGCTGTCCAGAAAGTGTGTGGTGCCACAAAGTGTGGGGGCTGGAGCCT	28374

[illegible]

TITLE	JOURNAL	REFERENCE	AUTHORS
Shanon,S., Naryanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schamman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.	Direct Submission	Submitted (26-FEB-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	3 (bases 1 to 196080)
Evans,G.A., Athanasou,M., Aguayo,P., Armstrong,D., Basil,M., Baetjer,J., Butler,C., Card,P., desailboat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schagman,J., Schultz,R.A., Stimson,S., Waller,K. and Ward,T.	Direct Submission	Submitted (24-FEB-1999) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA	On Feb 24, 1999 this sequence version replaced gi:2911733.
IMPOIRANT: This submission contains the entire insert of clone pD519013. pD519013 comes from the RPl-3 PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST AHAK. This region spans over 1.5 MbP. MARKER CONFIRMATION: EST: FTH (ferritin heavy chain mRNA), STS D11S659 and WI-7524			
MAPPED CLONE OVERLAP: PACs pD1466a11 and pD1756b9.			
FEATURES	SOURCE	Location/Qualifiers	
repeat_region	1..196080	/organism="Homo sapiens"	
repeat_region	/db_xref="taxon:9606"		
repeat_region	complement(2210..2482)		
repeat_region	/rpt_family="Alu"		
repeat_region	3457..3743		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(20145..20758)		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(23194..23599)		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(27367..27685)		
repeat_region	/rpt_family="Alu"		
repeat_region	28047..28457		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(29390..29565)		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(31076..31255)		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(34948..35045)		
repeat_region	/rpt_family="Alu"		
repeat_region	35232..35338		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(36242..36540)		
repeat_region	/rpt_family="Alu"		
repeat_region	37238..37522		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(38555..38650)		
repeat_region	/rpt_family="Alu"		
repeat_region	complement(40483..41164)		
repeat_region	/rpt_family="Alu"		
repeat_region	42042..42116		
repeat_region	/rpt_family="MIR"		
repeat_region	42677..42817		
repeat_region	/rpt_family="LRR12"		
repeat_region	42856..43157		
repeat_region	/rpt_family="Alu"		
repeat_region	43148..43380		
repeat_region	/rpt_family="LRR12"		
repeat_region	complement(43734..43999)		
repeat_region	/rpt_family="LRR12"		
repeat_region	complement(46998..47169)		
repeat_region	/rpt_family="MIR"		
repeat_region	complement(48048..48346)		

QY	82	-----	82	-----	82
Db	184653	CTTCCAGAGAGGCTGGGGGCTAGAGCCCGCTCGCAGCAGAAAGCTGGAGAGCCGAGCAT	184712		
QY	82	-----	82	-----	82
Db	184713	CGCGGGGCGCTGGGCGCTGGGCTCTGCGCCGACCTGGCCCTCGCCCTCGCCCGCG	184772		
QY	83	-----	83	-----	83
Db	184773	CCCTCCTGCCAGAGCTTCTACGTGAGCGTGTGTGTGACCCGCTGGTGAACCGATCGAG	184832		
QY	99	AsnLeuProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyAsp	118		
Db	184833	ACCTCGCGCGGCGCCACCGCCTCATGAGCTGTGTGTGGCTTCTCCAGAGCAGGAG	184892		
QY	119	GluGluGlyArgLeuLeuArgThrLeuLeuArgTyrAlaAsnLeuGlyAsnValLeu	138		
Db	184893	GAGCAGAGCGCGGCTGTGGGGGCGCAGCGTCATCGCTAGCCCAACTGGGCACGTCCTC	184952		
QY	139	IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal	158		
Db	184953	ATCTCTCGAGCGCTCGACACCGCAGCTCAAGCGCTTCCACAGCCCGCAGACCGTGTG	185012		
QY	159	Gln	159	-----	159
Db	185013	CAAGCAGGTGGGCGCGACCCGGGAGACAAAGGGGAGGCCACCGGGCAGAGCCAGGGCGCGAGAT	185072		
QY	159	-----	159	-----	159
Db	185073	GGGCGCGCAGGAATGGAAGATGGGTGAGCCAAAGTCCCGCGAGCTCGGGGATTTGGGT	185132		
QY	159	-----	159	-----	159
Db	185133	GGAGCGAGAGATGGGGGTGTGTGTCAAGATTTGGGGGTCCATTTGGGGGAGCAGAGTCCGG	185192		
QY	159	-----	159	-----	159
Db	185193	TGCTGAAGTGGGGCGAGGCCAGAGCCACCTCCGAGATGAGAGTCAAGCTGAGCAAGG	185252		
QY	159	-----	159	-----	159
Db	185253	CTAAGAGACCTTGAAGGATTAATGAAAGAGGCTAGCGGCTTGGAACTGTGAGTACT	185312		
QY	159	-----	159	-----	159
Db	185313	AGGCTACTTCCCTGTGCGCCTTGGCCCTCTTGATCTCCGGTTTCCACTCTGAGAGTATG	185372		
QY	159	-----	159	-----	159
Db	185373	GGACATTGGTCTGTACACACCCCTCAGCCTGGCTGACCTGGTCTGTGTTAATAAGACAG	185432		
QY	159	-----	159	-----	159
Db	185433	ACCCAGGCTAGGCGTGTGGCTCTCGCTGTAATCCAGTGTCTTTAGAGGCAAAAGTGG	185492		
QY	159	-----	159	-----	159
Db	185493	GAAATTCGCTTGAGCCAGCTGTTGAGACGCCCTGAGCAACATAGCGAGACCCCATC	185552		
QY	159	-----	159	-----	159
Db	185553	TCTACAAAACATTAAAAATTAGCAGGCGCATGTGGGCTGTGCTGTGATGTGAGGCTGA	185612		
QY	159	-----	159	-----	159
Db	185613	GTAATCGGAGGCTGAGCGAGGAGATCACTTGAACCCAGACAGTTCAGAGCTGAGTGCGC	185672		
QY	159	-----	159	-----	159
Db	185673	TAAATGCGCACCGCTGCACCTCAACCTTCGGTGACAGAGCCAGACCTTTCTCTGGAAATA	185732		

[illegible]

Db	187893	CCCTCCTATTATTTTGGTAATGGGGGTGAAGTCTGTGCTGCCCTTCTGTCACTG	187952
QY	357	-----	357
Db	187953	TGACACACACACACACACACACACACACACACACACACACATTCATTCCTG	188012
QY	357	-----	357
Db	188013	AAATTTCCCTGCCCCCAGTTATCTTTGTTTCCAGATCAAAACAAATCACACTTT	188072
QY	357	-----	357
Db	188073	TATGCTTGAATTTCTCCAGGGTGCCCGACGTGGCTCAAGATGTCCCTGGACCCCTAAG	188132
QY	357	-----	357
Db	188133	GCAGACCGGTGTACACTCTTCGGGGCTTTGTAGGGCATTTTAGAGTTGCTATTCACAGA	188192
QY	357	-----	357
Db	188193	ATCTGCCACCTAGACTGCCCCCTTATGTTCAAGCCACGTTCAATATATCTCTGTGCAT	188252
QY	357	-----	357
Db	188253	GAATGATAAATAATTATGCAACTCCAGGTAGATACATGAGGTGATAAAGGACGTACT	188312
QY	357	-----	357
Db	188313	CAGCCGAGTGAATACACTCAGGGACACTGTGGGTGTTCAAGGAGACCTGGCTCAGAMA	188372
QY	357	-----	357
Db	188373	GTTAGAGGGGCTGTGTCCAGAAGTGTGGTGGTCCCAAGTGTGGGGGGCTGGACCCT	188432
QY	357	-----	357
Db	188433	AAATCTGCCTTTGAAAGACAGTGTCCAGGACGAAAGGCTTCATGGGTGTGGAATAGC	188492
QY	357	-----	357
Db	188493	AGCAGCTGAGGTTTAAAGGGGGAAGCTGGCTTTGAGAGTTTGCCTGAGGGTTTACAGA	188552
QY	358	-----TTPGInSerSerSerSerThrProLeuGluArgMetMetMet 371	188553
Db	188553	GCCTCACCTGTCCCCAGAGTGGCAGAGCAGCTCAACACCCCTTGGAGAGATATATAT	188612
QY	371	tIleLeuArgProThrArgIleLeuSerThrGlyIleCys----- 383	188613
Db	188613	GATTTTGAACCACTGATTTGTGCACAGGAAATTTGCAAGTATGAGGAGAGAGAGAAA	188672
QY	383	-----	383
Db	188673	CCATACCATTGACCTTCCCAAAAGTGACCCAAAGAGAGAGACCCTACTTTCTTAGGGA	188732
QY	383	-----	383
Db	188733	GGCCTCAGATGATATATCAACCTTCCCTCCTCCTCCTGAGAGCAGTCAATTCACCTCA	188792
QY	383	-----	383
Db	188793	CAGATTTCTACCTCAATCTTTGAGGCTGACAGGACAGCACCATCTCCCATTTTCACAG	188852
QY	383	-----	383
Db	188853	CAGGGAACCTGAGGTCCAGAGAGAGAGAGATTCCTCAAGTCATCAGGCACATTCAAAG	188912
QY	384	-----	384
Db	188913	GTCTCGCTGGGATGATCTTTCTGTGGACCTTCTGTGCTCCCTGTGACACAGGTGCCCT	188972
QY	387	-----ArgCysProC 387	188973
Db	188973	ysTTPLeuTTPMetArgCysThrArgThrCysLeuGlyTTPSerArgThrCysThrGlyI	189032
QY	407	----- 407	189033
Db	189033	GTGTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	189092
QY	407	-----	407

OY 407 IeserProserHisSerProProthGlnLeuLeuProProSerValGluProProL 427
DB 189033 TATGCGCCGAGCCACACCCCTCCTCCGCGACATTCCTCGTGCAGCCCTCT 189092
OY 427 eutPrAlaProProSerThuser 434
DB 189093 TATGCGCTCCACCTTCACATCA 189115
RESULT 9
AC084857 133683 bp DNA linear HTG 22-NOV-2000
LOCUS AC084857
DEFINITION Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT
AC084857
AC084857
AC084857.1 GI:11276215
HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abram, H., Allen, N.,
Boutgen, S., Barre, N., Bastien, Y., Bede, F., Boguslavsky, L.,
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
DeRubeis, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Haefford, A., Horton, L.,
Ilav, I., Johnson, R., Jones, C., Kan, L., Karats, A., Labrecque, R.,
Lamarez, R., Landers, T., Lehotzky, J., Levine, R., Liu, C., Liu, G.,
Macdonald, P., Margus, N., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
Pierle, N., Pisan, C., Pollara, S., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schaner, S., Severy, P.,
Sougnier, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Trevers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.
DIRECT SUBMISSION
Submitted (22-NOV-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L11038
Center Clone name: 3231.N.5
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 126761 bases at least Q40
Consensus quality: 129952 bases at least Q30
Consensus quality: 131351 bases at least Q20
Insert size: 12800; agarose-fp
Insert size: 132483; sum-of-contigs
Quality coverage: 6.1 in Q20 bases; agarose-fp
Quality coverage: 5.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1459: contig of 1459 bp in length.
* 1460 1559: gap of 100 bp
* 1560 2520: contig of 961 bp in length
* 2521 2620: gap of 100 bp
* 2621 6334: contig of 3714 bp in length
* 6335 6434: gap of 100 bp
* 6435 11913: contig of 5479 bp in length
* 11914 12013: gap of 100 bp
* 12014 20756: contig of 8743 bp in length
* 20757 20856: gap of 100 bp
* 20857 29735: contig of 8879 bp in length
* 29736 29835: gap of 100 bp
* 29836 38645: contig of 8810 bp in length
* 38646 38745: gap of 100 bp
* 38746 52193: contig of 13448 bp in length
* 52194 52293: gap of 100 bp
* 52294 65209: contig of 12916 bp in length
* 65210 65309: gap of 100 bp
* 65310 79796: contig of 14487 bp in length
* 79797 79896: gap of 100 bp
* 79897 95896: contig of 16000 bp in length
* 95897 95996: gap of 100 bp
* 95997 113419: contig of 17423 bp in length
* 113420 113519: gap of 100 bp
* 113520 133683: contig of 20164 bp in length.

Location/Qualifiers
1. 133683
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CMD-3231N5"
/clone_id="CITD Human BAC"
1. 1459
/note="assembly-fragment"
/note="assembly-fragment"
1560. 2520
/note="assembly-fragment"
2621. 6334
/note="assembly-fragment"
6435. 11913
/note="assembly-fragment"
12014. 20756
/note="assembly-fragment"
20857. 29735
/note="assembly-fragment"
29836. 38645
/note="assembly-fragment"
38746. 52193
/note="assembly-fragment"
52294. 65209
/note="assembly-fragment"
65310. 79796
/note="assembly-fragment"
79897. 95896
/note="assembly-fragment"
95997. 113419
/note="assembly-fragment"
113520. 133683
/note="assembly-fragment"

BASE COUNT 31520 a 34761 c 35336 g 30857 t 1209 others
ORIGIN

Alignment Scores:
Pred. No.: 3, 13e-69 length: 133683
Score: 1278.00 matches: 381
Percent Similarity: 23.068 conservative: 0

Best Local Similarity: 23.06%
Query Match: 54.38%
DB: 2
Mismatches: 3
Indels: 1269
Gaps: 6
US-09-622-964-5 (1-435) x AC084857 (1-133683)
QY 51 ArgLeuAlaLeuThrGluGluGlnGlnLeuMetPheGluIuLysLeuThrLeuTyrCysAsp 70
DB 85158 AGGCTGGCCCTCAGCAGGAAGAACACACCTGATGTTTGAGAACTGACTGTATTTGGCAG 85217
QY 71 SerTyrIleGlnLeuIleProIleSerPheValLeu 82
DB 85218 AGCTACATACAGCTCATCCCATTTCTCTGCTGGGTGAGTTCCCTTCTGGCTGT 85277
QY 82 82
DB 85278 CCGGGTCCCTGTGCGCCGCCAGCTCCAGACAGGCGAGGAGAGTACAGAGAGCTGC 85337
QY 82 82
DB 85338 GCAAGGGCTGGGAGAGGGGGGAGCCAGCCAGGTCGGCCTCTCTGTAGG 85397
QY 82 82
DB 85398 AAGGTGCGGACTGCAGCCAGAACTGAGTTAGAGTTAGTAAGACTCTGCGT 85457
QY 82 82
DB 85458 TAGCAATGAAACCCCATTTTCTGAGGAGCGGTGACATGATGTCCTGGAGCCCTG 85517
QY 82 82
DB 85518 GCGGGAGAGGAGGGGTCTGGCGGATTCTGGGACACAGAGGGGACCCCGGGTGACA 85577
QY 82 82
DB 85578 GAACCTTTGGGCTCTCGCGCTCCATGAGGCTCTGCTGCTCTGCCAGAGCC 85637
QY 82 82
DB 85638 CTTCCAGAGAGGCTGGGGGCTAGGCCGCTCCGACGAAAGCTGAGAGCCGAGCAT 85697
QY 82 82
DB 85698 CGCGGGGCGCTGGGCTCTGGCGGAGCTGCGCCCTCGCCCTCGCCCGCCG 85757
QY 83 82
DB 85758 CCTCTGCGCCAGGCTTCTACGTGACGTGTGTGACCCGCTGGTGAACAGTACAG 85817
QY 99 AsnLeuProTPrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAsp 118
DB 85818 AACCTGCGGTGGGCGGACCGCCCATGACCTGTGTGCGGCTTCTCCAAAGGCAAGAC 85877
QY 119 GluGlnGlyArgLeuLeuArgTyrThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeu 138
DB 85878 GAGCAAGGCGGCTGCTGGCGGCGACGCTCATCCGTACGCCACCTGGGCAACGCTTC 85937
QY 139 IleLeuArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuVal 158
DB 85938 ATCCGTGCCAGGCTGACGACCGCAGTTCACAAAGCCCTTCCCAAGGCCACGACCTG 85997
QY 159 Gln 159
DB 85998 CAAGCAGGTGGGCGGACCGGAGCAAGCGGAGGACCGGCGACAGAGGCGCGAGAT 86057
QY 159 159
DB 86058 GGGGGCGGCAAGATGCAAGATGGTGGAGCCAAAGTCCCGGAGCTCGGGGGAATTGGGT 86117
QY 159 159
DB 86118 GAGGCCAGAGTGGGTGTGTCAAGATTGGGGGTCCAATTGGGCGGAGAGAGTGGG 86177

QY 159 159
DB 86178 TGTCTGAAGGTGGGGCGAGGCCAGAGGCCACCTCCGAGAGTAGAGCTGAGGACAGG 86237
QY 159 159
DB 86238 CTAAGACCTTGAGGGATATGGAAGAAAGGGGTGAGGGCTGGGATCTGTGAGTACT 86297
QY 159 159
DB 86298 AGGGTCTACTTCCCTCTGCGCTTCCCTTGTATCTCCGGTTTCCACTCTGAGGATG 86357
QY 159 159
DB 86358 GAGCATTTGCTCTGACACCCCTCAGCCCTGAGCTGACCTGTCTGTTAATAGACAG 86417
QY 159 159
DB 86418 ACCCAGCTAGCGGTGTGTGCTCTGCTGTATATCCAGTGTCTTGAAGGCAAGGTGG 86477
QY 159 159
DB 86478 GAAATCGCTTGAAGCCAGCTTTGAGAGCGCCCTGAGCAACATAGCGAACCCCATC 86537
QY 159 159
DB 86538 TCTCAAAACATTAATAATTAGCAGGGCATGTGGCTGTGCTGTAGTGTAGGCTGA 86597
QY 159 159
DB 86598 GTATCGGAGGCTGAGGAGGAGATCACTTGAAGCCAGAGTTCAGAGCTGCACTGCGC 86657
QY 159 159
DB 86658 TAAGATGCACCGCTGCATCTCAACTCGGTGACAGAGCCAGACCTTCTGTGAATA 86717
QY 159 159
DB 86718 AATTAATACCTGCGCACATGCTCAGCCAGAACAGCACTAGAGTGTCTCAAAATT 86777
QY 159 159
DB 86778 TTTTGTGTTGAAAGAAAGAGATGGCAAGAGTGTGAGTTCCTATAGTCAAGAG 86837
QY 160 160
DB 86838 TGCGGGCATCCCTTCTGAGGTTCTCCACCCACCTGCTTCTTCACTCCACTCTGAGG 86897
QY 161 yPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsnMetPr 181
DB 86898 CTTTATAGACTCGGCGAGAACACAGCAGTGGAGAACTGAGCTTACCAACACATGTT 86957
QY 181 eTyrValProTPrValTPrPheAlaAsnLeuSerMetLysAlaTPrLeuGlyArgTyr 201
DB 86958 CTGGGTGCGCTGGGTGTGTGTTGGCCAACTGTCAATGAAGGCGTGGCTTGGAGTGTG 87017
QY 201 eArgAspProIleLeuLeuGlnSerLeuLeuAsn 212
DB 87018 CCGGGACCTTATCTGCTGCTCAGAGGCTCTGTGAACGTGAGCCCATCTAGACAGAGGCTG 87077
QY 212 212
DB 87078 CCGCAGAGTGGAGAGGTTGTGTCCACAGAAACAGGTTTCTACAAAGAGACCTT 87137
QY 212 212
DB 87138 GGGGCCCTGAGGGTCTTCCAGAGACCTGAGGTGGGTTGCAAACTTTTCCAAACAGCA 87197
QY 212 212
DB 87198 TCACAGCCCGAGGTGTGCTTCTCTAGAGGCCCTCTCTCTTCTCCAAAGTCTGTAGGT 87257

QY 212 ----- 212
Db 87258 CCTGGTCCCTTTGATAGATGAGAAAGCTGAGACAAAGAGCTTTAGTGAGCTTCCCA 87317
QY 212 ----- 212
Db 87318 TGGCCACAGACCCAGGAATGACCATAGTACAGGCCCTGTACTGTGAGAAAGAGTGG 87377
QY 212 ----- 212
Db 87378 GGGGAGACCCAGGCTGGGGGAGGTGGTGTTCAGAACCCATCCCTCTCTGCCCCC 87437
QY 213 --gluMetasnThrLeuArgThrGlnCysGlyHisLeuTyrAlaTyrAspTrpIleSerI 232
Db 87438 AGGAGATGAACACCTTCGCTACTCAGTGTGACACCTGTATGCTTACGACTGATTA 87497
QY 232 LeProLeuValItyrThrGlnVal----- 239
Db 87498 TCCCACTGGTGTATACACAGGT--GAGGACTAGGCTGGTGAAGCTGCCCTTTGGAAACT 87556
QY 239 ----- 239
Db 87557 GAGGCTAGAAAGAACAAAGAACAGCTGGGGTGGAAAGGCTCACTAGAGCTTAAGTGG 87616
QY 239 ----- 239
Db 87617 CTCCCTGGGAGTGGGTCCACACTTTGAAGTTGGGTCTGAGCTTTGAAGTGCAGTTC 87676
QY 239 ----- 239
Db 87677 TAAAGATCCAGGCTCCTGCTGCGCCAGTCCAGTAGAGCAATGTATATCCCAATTT 87736
QY 239 ----- 239
Db 87737 AAAGAGAGTGTGGCCGGGACACAGTGGCTATGCTGTATCCAGCACTTGGGAAGCTG 87796
QY 239 ----- 239
Db 87797 AGGCAAGGTGATCACTGAGTGCAGAGTTCGAGACAGCCTGGCCAACTAGTGAATCC 87856
QY 239 ----- 239
Db 87857 CCATCTCTACTGAATAATACGAATTAGCTGTGTGTGTGTCACAGCCTGTATATCCAGCTA 87916
QY 239 ----- 239
Db 87917 CTGGGAGGCTGAGCAGAGAAATCGCTTGAAACCCGGGAGTGGAGGTTGCATGAGCTG 87976
QY 239 ----- 239
Db 87977 AGATCATGCCACTGCACCTCCAGCCTGGGCGACACAGCAAGACTGTGTCTCAACAAACA 88036
QY 239 ----- 239
Db 88037 ACNAACAAACAAACAAACAAAGGGTTAACAGAGCCCTTAAGTCACTAAGTGTGC 88096
QY 239 ----- 239
Db 88097 AAGTCAGAACAGGCTGTGTCTCTGTCTCAGACTCCAGCCCTGAGACATCTGATTT 88156
QY 239 ----- 239
Db 88157 TCAGGTTGCCACCTAGCCCTTTGCTACACATCTCTCTCTCTCTCTCTCTCTCTCCAG 88216
QY 240 --ValThrValAlaValItyrSerPheLeuThrCysLeuValGlyArgGlnPheLeu 259
Db 88217 TGGTACTGTGGGGGTGTACAGCTTCTCTGACTGTCTAGTGGGCGGAGTTTCTCA 88276
QY 259 snProAlaIysAlaItyrProGlyHisGlnLeuAspLeuValItyrProValPheThrPheLeu 279
Db 88277 ACCCAGCCAAAGGCTTACCTGTGGCATGAGACTGTGACCTGTGGCCGCTTCAACGTTCC 88336
QY 279 euGlnPhePhePhePheTyrValGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeuGlyT 299

Db 88337 TGCAGTCTCTCTCTATATGTGGCTGTGGTGAAGGGGCGCTCTCCAGGGCCCTGCTGGCT 88396
QY 299 rParGHisGlyGlnArgGlyHisGlyGlnGlnLeuLeuGluThrArgMetGlnCysGlnG 319
Db 88397 GAGGCAATGGCCAGAGGGGTGATGGCCAGCAAGCTGCTGTAGACGAGATGATGATGACG 88456
QY 319 IuArgLysValIserArgValGluSerSerGlnAlaTrpTrpArgThrProValIlePro 339
Db 88457 AAAGGAAGTCTCAGCGGTAGAAAGCAGCAGCGGTGTGGCACACACTGTATATCCAG 88516
QY 339 IatThrArgGluAlaGlnAlaGlyGluSerLeuGluProGlyArgArgArgLeuTrp 357
Db 88517 CTACTGGGAGGCTGAGGCGAGAGAAATCGTTGAACCCGGGAGGCGGAGGTTGTGTGAG 88576
QY 357 ----- 357
Db 88577 TTGAGATCGTGCACCTGCACTCCAGCCTGGGCAAAAGATGAATCTATCTCAAAAACA 88636
QY 357 ----- 357
Db 88637 ACAACAAACAACAAACAAAGCCCTAAGTTCAGAAAGCCCTTGCCCTTTAGAGGAGAGC 88696
QY 357 ----- 357
Db 88697 GGCACCACTCTCTTATTTCAAGATGCTGTGGGCTGTCTGTCTCTCACTCAAGTGG 88756
QY 357 ----- 357
Db 88757 CTGTCCAGTATTTCCCTCCACACAGCAATACCTCCGAACAGATGTTCTGAATCAGAC 88816
QY 357 ----- 357
Db 88817 AGTTTCTCTCCACTTTTATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88876
QY 357 ----- 357
Db 88877 TCTACTCTCTTATTTTGTGTAATGGGGGTGTAAAGTCTGTCTCTCTCTCTCTCT 88936
QY 357 ----- 357
Db 88937 CACTGTGACACACACACACACACACACACACACACACACACACATTCCTATTC 88996
QY 357 ----- 357
Db 88997 TCTAATTTCCCTCGCCGCCCGAGTTATCTTGTGTTCTGTGAGATCAACAAATCAGAC 89056
QY 357 ----- 357
Db 89057 TTTTATCTTGAATTTCTCAGAGTCCCGCAGTGGCCTCAAGATGTTCCCTGGACCT 89116
QY 357 ----- 357
Db 89117 AAGGACAGCGGTGCACCTCTTCGGGGCTTTGTAGGCAATTTAGAGTTGCTATCA 89176
QY 357 ----- 357
Db 89177 GGAATCTGCCAAGCTAGACTGCCCTTTACTTCAAGCCCACTTCAGTATATCTCTGTG 89236
QY 357 ----- 357
Db 89237 CATGAATGAATTAATTAATTAAGCACTCCAGTAAGATACATGAGTGAATAAAGCAAGTG 89296
QY 357 ----- 357
Db 89297 ACTCAGCCAGTGTATCACTCAGGACAGCTGTGGGTTCAGGGAAGAGCTGGCTCAGA 89356
QY 357 ----- 357
Db 89357 AGAGTTAGAGGGGCTGTGTCCAGAAAGTGTGTGGGTGCCACAAGTGTGGGGGCTGGAGC 89416
QY 357 ----- 357


```

source
1. .160169
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-810P12"
1. .1926
/misc_feature
/note="assembly_name:Contig18"
2027. .4202
/note="assembly_name:Contig19"
4303. .7163
/note="assembly_name:Contig20"
clone_end:T7
vector_side:left"
7264. .9961
/misc_feature
/note="assembly_name:Contig21"
10062. .12575
/note="assembly_name:Contig22"
12676. .15545
/note="assembly_name:Contig23"
15646. .20941
/note="assembly_name:Contig24"
21042. .23966
/note="assembly_name:Contig25"
24067. .27231
/note="assembly_name:Contig26"
27332. .30550
/note="assembly_name:Contig27"
30651. .35202
/note="assembly_name:Contig28"
35303. .40201
/note="assembly_name:Contig29"
40302. .45758
/note="assembly_name:Contig30"
45859. .51249
/note="assembly_name:Contig31"
51350. .58267
/note="assembly_name:Contig32"
58368. .66916
/note="assembly_name:Contig33"
67017. .77145
/note="assembly_name:Contig34"
77246. .86493
/note="assembly_name:Contig35"
86594. .96688
/note="assembly_name:Contig36"
96789. .107169
/note="assembly_name:Contig37"
107270. .117047
/note="assembly_name:Contig38"
117148. .127120
/note="assembly_name:Contig39"
clone_end:SP6
vector_side:right"
127221. .142029
/note="assembly_name:Contig40"
142130. .160169
/note="assembly_name:Contig41"
BASE COUNT 41647 a 36693 c 37296 g 42226 t 2307 others
ORIGIN
Alignment Scores:
Pred. No.: 1.53e-67 Length: 160169
Score: 1252.50 Matches: 378
Percent Similarity: 22.96% Conservative: 0
Best Local Similarity: 22.96% Mismatches: 6
Query Match: 53.30% Indels: 1266
DB: 2 Gaps: 6
US-09-622-964-5 (1-435) x AC051664 (1-160169)
OY 51 ArgleuAaLeuThrGluGluGlnGlnLeuMetPheCyluLysLeuThrLeuTyrCysasp 70
|||||
74735 AGGCTGACCTTCACGGAGAACACAGCTCATATTTTGAGAACACTGCTGATATTCGCAC 74676
|||||

```

OY	71	sercrrllleqlleuilepilleserPhavalLeu-----	82
Db	74675	AG-TACATCCAGCTCATCCCATTTCTCTGTGCTGGTGAATTCCCTTGTGCTGTT	74617
OY	82	-----	82
Db	74616	CCGGGTCCCTGTGGCCGCCAGGCTCCAGACAGGCGCAGAGGAGATCAGAGAGCTGGC	74557
OY	82	-----	82
Db	74556	GCAAGGGGCTGGGAGGGGGGGGGAACGACAGCGAGAGGTGCGCCTCTGTAGGA	74497
OY	82	-----	82
Db	74496	AAGGTGGGAGCTGACAGCCATGAAACTGAAGTTAGACGTTAGSTAAAGAGTCTCGCCTT	74437
OY	82	-----	82
Db	74436	AGCAATGMAAACCCCATTTTCTGAGGGAAAGCGCTGACATCATGTCCTTGAGCCCTGC	74377
OY	82	-----	82
Db	74376	GCGGAGGGAGGGGGGCTGTGGCGATTTCTGGGACCAAGAGGGGGACCCGGGTGACAA	74317
OY	82	-----	82
Db	74316	ACCCTTGGGGCTCTCGGCGCTCCATGCGAGGCTCTGCTGCTCTGCCAGCGCT	74257
OY	82	-----	82
Db	74256	TCCAGAGGGCTGGGGGCTAGGCCCGCTGCGACGAAGCTGAGAGCCGAGGATCG	74197
OY	82	-----	82
Db	74196	CCGGCGTGGGCCCTGGGCTCTGGCGCGAGCGTGGGCCCTCGGCCCTCCGCCGCCGCC	74137
OY	83	-----GlyPheTyrValThrLeuValValThrArgTPTripasnGlnTyrGluasn	99
Db	74136	TCCGTGCCAGCGCTTACAGTGAGCGGTGCTGTCGCCGCTGGTGGGAACAGTACGAGAAC	74077
OY	100	LeuProTTPProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGlu	119
Db	74076	CTGCCGTGGCCGACCGCTCATGAGCGCTGTGCTGGGCTGCTCGAAGGCAGGCGAG	74017
OY	120	GlnGlyArgLeuLeuArgTrrThrLeuIleArgTyrAlaasnLeuGlyLysValLeuIle	139
Db	74016	CAAGCGCGGCTGTGGGGGCGACAGCTCATCCGCTACGCCAACCTGGGCAACGTCATC	73957
OY	140	LeuArgSerValSerThrAlaValIlyTyrAsnArgPheProSerAlaGlnHisLeuValGln	159
Db	73956	CTCGCGAGCTCTAGCACCGCAGTCTCAAAAGCGCTTCCACGCGCCACGACCTGTGCAA	73897
OY	160	Ala-----	160
Db	73896	GC-AGTGGGCGGACCGGGAGCAAGGGGAGGCGACCGGGCAGAGCCAGGGGGCGAGATGG	73838
OY	160	-----	160
Db	73837	GCGCGCAGAGATGAAGATGGGTGGAGCCAAAGTCCCGGACTCGGGGATTTGGGTG	73778
OY	160	-----	160
Db	73777	GAGCCAGAGAGTGGGGTGTAGTCAAGATTTGGGGGTCCAAATTGGGGCGGGACAGATCGGCT	73718
OY	160	-----	160
Db	73717	GTCCTAAGAGTGGGCGAGGCGACAGAGCCACCTCCGAGATGAGAGTCTGAGCGAGGC	73658
OY	160	-----	160
Db	73657	TAAAGACCTTGAGGGATTAATGMAAGAAAGGTGACGCGCTTGGAACCTGGTAGTACTA	73598

```
OY 160 ----- 160
Db 73597 GGGTCTACTCCCTCGCCCTTGCCTCTTGATCTCCGGTTCCACTCTGAGGTATGG 73538
OY 160 ----- 160
Db 73537 GACANTTGGTCTGTGACACCCCTCAGCCTTGCGCTGACCTGTGCTCTGTTAATAAGACAGA 73478
OY 160 ----- 160
Db 73477 CCCAGGCTAGGCGTGGTGGCTCTCGCTGTATATCCAGTGCCTTAGAGGCAAAAGTGGG 73418
OY 160 ----- 160
Db 73417 AAGATCGCTTGAGCCAGCTGTTTGAGACGCCCTTGAGCACAATAGCAGACCCCATCT 73358
OY 160 ----- 160
Db 73357 CTACAAAAACATTAAAAATTAGCAGGCGATGGTGGCGTGGCTGCACTGAGGTGAG 73298
OY 160 ----- 160
Db 73297 TATGGGAGGCTGAGGCGACGAGAGATCACTTGAGCCAGAGTTCAGCGCTGCAGTGGCT 73238
OY 160 ----- 160
Db 73237 AAGATCCGACCGCTGCACTCCAACTGGTGAAGGCCAGACCCCTTCTCTGAAATTA 73178
OY 160 ----- 160
Db 73177 ATAAATACCTGCCACATGCTCAGCCAGAACAGCACCTAGTAGTGTCTGAAATTTT 73118
OY 160 ----- 160
Db 73117 TTGTGTGTGAAGAAAGAGAGATGCAAGAGTGTGAGTGTCTTATAGTTCAGCAGGT 73058
OY 161 ----- G1Y 161
Db 73057 GCCGCCATCCCTTCGACAGTTCTCCACCCACCCCTTCTTACCTCCACTCTGCAGGC 72998
OY 162 PheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMetPhe 181
Db 72997 TTTATACCTCCGCGACAAACAGACAGTGTGAGAACTGAGCCCTACACACAAATGTTTC 72938
OY 182 TrpValProThrValIleThrPheAlaAsnLeuSerMetLysAlaTrpLeuGlyLysArgIle 201
Db 72937 TGGGTCCCTGGGTGGTGGTTTGGCACTGTCAATGAAAGCGTGGCTTGGAGGTGCAATC 72878
OY 202 ArgAspProIleLeuLeuGlnSerLeuLeuAsn ----- 212
Db 72877 CGGGACCTATCTCTGCTCCAGAGCCTGTGACGTGAGCCACTGTACAGACAGGGCTGC 72818
OY 212 ----- 212
Db 72817 CGCAGAGTGGAAAGGCTGTGTCCACAGGAAACAAGTTTCTTACAAAGAGAACCTTGG 72758
OY 212 ----- 212
Db 72757 GGCCTTGAGGGTCTTCCGAGAGCCTGAGGTGGGGTTGCAGAAATCTTTTCCACAGCAATC 72698
OY 212 ----- 212
Db 72697 CACAGCCCGAGGTGGTCCCTTCTCAGAGGCCCTCCCTTCTTCCAAAGTCTGTAGAGTCC 72638
OY 212 ----- 212
Db 72637 TGGTCCCTTTTGATAGATGAGAGAGCTGAGACACAAAGGTTTAGTGAAGTTCCCATG 72578
OY 212 ----- 212
Db 72577 GCCACACAGCAGGAATGACCATAGTACAGAGCCCTGTGTACTGTGAGAAAGAGTGGGG 72518
OY 212 ----- 212

OY 72517 GCGAGCCGAGGTTGGGGGACAGTGTGTTTCAGAACCCCATCCCTCTTTCGCCCCAG 72458
OY 213 GluMetAsnThrLeuAsnArgThrGlnCysGluHisLeuTrpAlaTrpAspTrpIleSerIle 232
Db 72457 GAGATGAACACCTTGGCTGCTACTCAGATGTGACACCTGTATGGCTACGACTGATTAATATC 72398
OY 233 ProLeuValIleThr ----- 237
Db 72397 CCACGTGTATACACAGGTGAGACTAGGCTGTGAGGCTGCCCTTTTGGGAAACTGAG 72338
OY 237 ----- 237
Db 72337 GCTAAGAGACCAAGAGACAGCTGGGGTGGGAAGGCTCACCTAGAGGCTAAGTGGCTC 72278
OY 237 ----- 237
Db 72277 CCTTGGAGTTGGGTCCACACTTTGAAGTTGGGTCTGAGCTTGAAGTGCACAGTTCTAA 72218
OY 237 ----- 237
Db 72217 GAGTCCAGGCTCTGCTGCTGCGCCAGTCCAGTAGAGCCAAATGTGATATCCCATATTAAA 72158
OY 237 ----- 237
Db 72157 GAGAGTTGGCGGGTGCAGTGGCTCATGCTGTATATCCAGCACTTTGGGAAGCTGAGG 72098
OY 237 ----- 237
Db 72097 CAGGTGATCACCCTGAGTCAAGAGTTCAGAGACAGCCTGGGCAACATGTGAAACCCCA 72038
OY 237 ----- 237
Db 72037 TCTTACTGAAAAATACAGATTAGCTGTGTGTGTGTCAGCGCTGTAAATCCAGTACTT 71978
OY 237 ----- 237
Db 71977 GGGAGCTGAGGACGAGAGATGCTTGAACCCGGGAGGTGAGTTCAGTGAAGTGA 71918
OY 237 ----- 237
Db 71917 TCATGCCACTGCACACTCCAGCCTGGGCGACACAGCAAGACTGTGTCTCAAAACAACAACA 71858
OY 237 ----- 237
Db 71857 AACAAACAACAAAGGGTTAAACAGAGCCCTTAAGTACATTAAGTGTGCAAGTCAAGAAC 71798
OY 237 ----- 237
Db 71797 AGGCTTGTGTCTCTGTCTCAGACTCCAGCCCTGAGACATCGAATTCAGGGTTCCC 71738
OY 238 ----- GlnValIleThrValIleValIleTrp 245
Db 71737 ACCTAGCCCTTGTGTACACATCTCTCCCTCCCTCCACAGGTGTGACTGTGGGTGAC 71678
OY 246 SerPheLeuLeuThrCysLeuValIleArgGlnPheLeuAsnProAlaLysAlaTrpPro 265
Db 71677 AGCTTCTTCTGACTGTGTAGTTGGGCGGAGATTTCTGAACCCGACCAAGGCTTACCT 71618
OY 266 GlyHisGlnLeuAsnProAlaValIleProValIlePheThrPheLeuGlnPhePheThrVal 285
Db 71617 GGGCATGAGCTGACCTGCTGTGTGCGCTTCCGTTCCAGTTCGCAAGTCTTCTCTCATGT 71558
OY 286 GlyTrpLeuLysValIleLysSerArgAlaLeuLeuGlyTrpArgHisGlyIleArgGly 305
Db 71557 GGCTGGCTGAAGGTGGGCTCTCCAGGCGCCCTGTGCGGTGAGGCAATGGCCACAGGGGT 71498
OY 306 HisGlyGlnGlnLeuLeuGlnThrArgMetGlnCysGlnIleLysValIleSerArgVal 325
Db 71497 CATGGCCACAGCTGCTGCTGAGCAGAGATGCACTGTACAGAAAGAAAGTGTCTCAGGGTA 71438
OY 326 GluSerSerGlnAlaIleTrpArgThrProValIleProAlaThrArgGluAlaGluAla 345
```

Db	71437	GAAGGACGCCAGGGGGTGGTGGCCACACACCTGTAAATCCAGCTACTGGGAGGCTGAGGCA	71378
OY	346	GLYIuseIserLeuGluP-roGlyAraGlyLeuTrp-----	357
Db	71377	GGAGAAATCGCTTGAAACCCGGGAGCGGAGGTTGTGTGAGTTGATGATCGTGCACCTGCAC	71318
OY	357	-----	357
Db	71317	TCCAGCTGGGGCAAAAAGATGAAGACTCTATCTCAAAAACACACACACAAACAAAG	71258
OY	357	-----	357
Db	71257	CCCTAAGGTTCCAGAGAGCCCTGCCCTTAAAGAGCAGATCGGCACCACCTCTCTATTC	71198
OY	357	-----	357
Db	71197	AAGATGCCCTGTGGGCTGTCTTTCTCTCACTCAAGTGGCTTGTCCAGGATTCCTCC	71138
OY	357	-----	357
Db	71137	CACCACAGCCAAATACCTCCGAAAGATGTTCTGAATCACAAGTTTCTCTCACCTCTT	71078
OY	357	-----	357
Db	71077	ATCTTTCCTTCCTTCTGTGTGCCCCACCACTCTCTCTCTCTACCTTCTTATTTT	71018
OY	357	-----	357
Db	71017	TGGTAATGGGGGGTGAAGTCTGTGTCTGTGCCCTTCTGTCACTGTGACACACACACA	70958
OY	357	-----	357
Db	70957	CACACACACACACACACATACACACACACACACACAGCATTCCTATTCCTAAATTC	70898
OY	357	-----	357
Db	70897	CCCGCCCCCCAGTTATCTTTGGTTTCTGTGCAGATCAAAACAAATCACACTTTATGTT	70838
OY	357	-----	357
Db	70837	GAATTTCTCCGGGGTGGCCCACTGGCCTGCAGAGATGTCCCTGAGACCCCTAAGCAGAG	70778
OY	357	-----	357
Db	70777	CGGTCACTCTTCCTGGGGCTTTTGTTAGGGCAATTTAGAGGTGTCTATCCAGAAATCTGCC	70718
OY	357	-----	357
Db	70717	CACCTAGACTGCCCTTATGTAGTTCAGCCAGCTTCAGATATATCTCTGTTGCATGATGA	70658
OY	357	-----	357
Db	70657	TAAAAATTATGCAACTCCAGGTAGATACATGAGTGAATAAAGCAGTACTCAGCCGA	70598
OY	357	-----	357
Db	70597	GTATATACACTCAGGAGACAGCTGTGGGTGTTCAGGAGAGACATGGCTCAGAAAGTTAGAG	70538
OY	357	-----	357
Db	70537	GGGCTGTGTCCAGAGAGTGTGTGGTCCACAAAGTGTGGGGCTGAGCCCTAAACTGTC	70478
OY	357	-----	357
Db	70477	CTTTGAAGACATGTGTCAAGCAGGAAGGGCGTCATGGGTGTGAAATATGACGACAGCTGA	70418
OY	357	-----	357
Db	70417	GGTTTAAAGGGGGAAGCTGGCTTTGAGAGATTCTGCTGAGGTTTACAGAGCTCACCT	70358
OY	358	-----TrGlnSerSerSerSerThrProLeuGluArgMetMetLileuAr 374	
Db	70357	GTCCCCAAGGTGGCAGACAGCTCATCAACCCCTTTTGGAGAGAGATGATCATATTTTGA	70298

QY	374	gprthnglyleuSerthnglylecyasarg	-----	384
Db	70297	ACCAATGATGATTGCGACGAAATTGCG- GATGGGAGAGGAGAAACATACCA		70233
QY	384	-----		384
Db	70238	TGGACCTTCCCAAAGTGAGCCAAAGAGAGAGACCCACTGTTCTGTAGGGAGGCTCAC		70179
QY	384	-----		384
Db	70178	AGTGAATGATCAACCTTCCCTCTCTCCCTCTGACAGCAGTCATTCACACAGATTC		70115
QY	384	-----		384
Db	70118	TCACCTCATCTTTGAGGCTGCGAGCGACGCCATCTCCCATTTCCACAGCGAGGAAA		70055
QY	384	-----		384
Db	70058	CTGAGGTCCAGAGAGAGGGAGATTCCTCCAGTCATGACGACATAGATGCTTGC		69999
QY	385	-----CysProCysTrrple		389
Db	69998	CTGGGATGATCTTCTGTGGGACTTCTCTGTCCTCTGTCGACGAGGTGCCCTGTGGCT		69933
QY	389	UTTPMAtArgCysAtArgthrcysleuglytrpSarArgthCysThnglyleSerP		409
Db	69938	GTCGATGATGATGACACCGACGCTGCTCGGATGAGCCGACATGTACTGGAAATAGCCC		69879
QY	409	oSerHlIsSerProProthrcGlnleuLeuProProSerValGluProProLeuTrpAl		429
Db	69878	GAGCAGACAGCCCCCTACACAGCTGT- CGGCCAGTCCGTCGACGCTCTTATGGGC		69820
QY	429	aProProSerThrcSer 434		
Db	69819	TCCACCTTCACATCA 69804		
RESULT 11				
LOCUS	AF440756			
DEFINITION	Homo sapiens vitelliform macular dystrophy 2-11ke protein 1 mRNA,			
ACCESSION	AF440756	1908 bp	mRNA	linear
VERSION	AF440756			
KEYWORDS	AF440756.1 GI:21734839			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	1 (bases 1 to 1908)			
REFERENCE	Marguardt, A., Stohr, H., Passmore, L.A., Kramer, F., Rivera, A. and			
AUTHORS	Weber, B.H.			
TITLE	Mutations in a novel gene, VMD2, encoding a protein of unknown			
REFERENCE	properties cause juvenile-onset vitelliform macular dystrophy			
AUTHORS	(Best's disease)			
TITLE	Hum. Mol. Genet. 7 (9), 1517-1525 (1998)			
REFERENCE	98367043			
AUTHORS	9700209.			
TITLE	2 (bases 1 to 1908)			
REFERENCE	Stohr, H., Marguardt, A., Nanda, I., Schmid, M. and Weber, B.H.			
AUTHORS	Three novel human VMD2-like genes are members of the evolutionary			
TITLE	highly conserved RFP-TM family			
REFERENCE	Eur. J. Hum. Genet. 10 (4), 281-284 (2002)			
AUTHORS	22027749			
TITLE	12032738			
REFERENCE	3 (bases 1 to 1908)			
AUTHORS	Stohr, H., Marguardt, A. and Weber, B.H.F.			
TITLE	Three novel human VMD2-like protein genes are members of the			
REFERENCE	evolutionary highly conserved RFP family			
AUTHORS	4 (bases 1 to 1908)			
TITLE	Unpublished			
REFERENCE	Stohr, H., Marguardt, A. and Weber, B.H.F.			
AUTHORS	Direct Submission			

[illegible]

PUBMED 9700209
 2 (bases 1 to 1506)
 Stohr, H., Marguardt, A., Nanda, I., Schmidt, M. and Weber, B.H.
 Three novel human VMD2-like genes are members of the evolutionary
 highly conserved RFP-TM family
 Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
 JOURNAL 22027749
 MEDLINE 12032738
 PUBMED 12032738
 3 (bases 1 to 1506)
 Stohr, H., Marguardt, A. and Weber, B.H.F.
 Three novel human VMD2-like protein genes are members of the
 evolutionary highly conserved RFP family
 Unpublished
 4 (bases 1 to 1506)
 Stohr, H., Marguardt, A. and Weber, B.H.F.
 Direct Submission
 Submitted (30-OCT-2001) Human Genetics, University of Wuerzburg,
 Biozentrum, Am Hubland, Wuerzburg 97074, Germany
 JOURNAL
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 SOURCE
 Location/Qualifiers
 1. 1506
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="12"
 /map="12q14.2-q15"
 173. 1369
 /note="VMD2L3: putative membrane protein; Pfam01062;
 worm_family=8"
 /codon_start=1
 /product="vitelliform macular dystrophy 2-like protein 3"
 /protein_id="AA076997.1"
 /db_xref="GI:21734844"
 /translation="MTVTYSSKVNATFEFGRHLILKNGSTLYKLLREIFYEALVLT
 AILSYRLTLGVOKRFEKISTYCDRAEOIPYEVGCGFYTLVYVPMWPMVLPW
 PDRMLFLSSVHSGDSEGRRLRLRLVNTSLITRSVTAHYKRPIDHVEA
 GFMTDERKLNLKSPHLKTYWPFIMFGNATKANEGRIDSVDLSTLMDENRYR
 SWCSLFGYDMVGIPLVYTOVTLAVYTFEFCILGRQFLDPTKGYAGLDLILYPIF
 TLQFFYAGMKVAEQLINPEGDDDEFETWICIDRLQVSLAVDEKHMSLPKMK
 DIYMDSSARPPYTLAADYCIPISEFISGVGKQMPKMKEMEDIKIPLPQPCQA
 KSDPGC"
 BASE COUNT 416 a 341 c 330 g 419 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.24e-58 Length: 1506
 Score: 1063.00 Matches: 217
 Percent Similarity: 60.00% Conservative: 44
 Best Local Similarity: 49.89% Mismatches: 101
 Query Match: 45.23% Indels: 73
 DB: 9 Gaps: 3
 US-09-622-964-5 (1-435) x AF440758 (1-1506)
 QY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 DB 173 ATGACTGTCACCTACTCCAGTAAGTACCAATGCAACTTTTTCGATTTCATAGCTTA 232
 QY 21 LeuLeuGlySTPARGlySerIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 233 CTTCCCAAGTGGAGGCGACATCTACCAACTACTGTACAGGAAATTTGTTTGGCT 292
 QY 41 LeuGlyTyrTyrIleIleArgPheIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
 DB 293 GTTCTTTACACGACATAGTTTGGTACAGATTCTTACCTACAGAGTCCAAAACGT 352
 QY 61 MetPheGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 DB 353 TACTTGAATAATATCAATTACTGTGACAGATATGTCGACCAAAATTCAGTAACCTTT 412
 QY 81 ValLeuGlyPheLeuValThrLeuValThrArgTyrTrpAsnGlnIleArgLeu 100
 DB 413 GTGCTGGGTGTTTATCTCTCTGTAAGACCGATGGTGGACCAAGTTTGAAATTGG 472
 QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluLeuLeuLeuLeu 120

DB 473 CCTGGCCGACAGAGGTATGTTCCATCTGTAGCAAGTGTTCACGGAAGCGACAGCAC 532
 QY 121 GlyArgLeuLeuArgArgThrLeuLeuArgTyrAlaAsnLeuGlnValLeuLeu 140
 DB 533 GGGCCGCTGCTTAGAAGCGGTGAGCCCTACGTCAATCTCACTCCCTGCTCATCTTT 592
 QY 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnIleLeuValGlnAla 160
 DB 593 CGCTGGTAGACACTGCTGTGTACAAAGATTTCCACATGACACACGCGTGTGAAGCA 652
 QY 161 GlyPheMetThrProAlaGlnIleLysGlnLeuGlnLysLeuSerLeuProIleAsnMet 180
 DB 653 GGTATTATACACACAGATGAAGAAATATTCAACACCTCAAGTCTCTCATCTGAAA 712
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
 DB 713 TATTGGGTTCCATTCTCTGTTGGAAATCTTGCAACTAAACCCGGAATGAAGTAGA 772
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnIleAsnThrLeuArgThrGln 220
 DB 773 ATCAGAGACAGTGTATCTGCATCTGATGATGACGAAATGAATGATACCGCTCTGG 832
 QY 221 CysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValAla 240
 DB 833 TGCAGCCTCTTATTCGGTATGACCTGGTGGATTCGCGTGTATACACCGAGTGTGC 892
 QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 DB 893 ACTCTGCTGTATACCTCTCTCTTCTGCTGCTTATGGACCGCAAGTTTGGATCCC 952
 QY 261 AlaLysAlaTyrProGlyHisGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
 DB 953 ACCAAGGCTACGACGAGGATGACTGATCTTACATTCATCCATCCATCCCTACCA 1012
 QY 281 PhePhePheTyrValGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeuGlyTrpArg 300
 DB 1013 TTCTTCTCTATGACAGATGAGCTTAAGTA- 1042
 QY 301 HisGlyGlnArgGlyHisGlyGlnLeuLeuGlnIleThrArgMetGlnCysGlnGluArg 320
 DB 1042 ----- 1042
 QY 321 LysValSerArgValGlnSerSerGlnAlaTrpTrpArgThrProValIleProAlaThr 340
 DB 1042 ----- 1042
 QY 341 ArgGluAlaGluAlaGlyGlnSerLeuGlnProGlyArgArgLeuTrpTrpGln-Ser 360
 DB 1043 -----GCAAG 1048
 QY 360 rSerSerSerThrProLeuGlnArgMetMetIleLeuArgProThrGlyLeuSerThr 380
 DB 1049 CACCTTATCAACCTTTTGGAGAAGATGAAGATGATTTTAAACTAAGTCTGTGATGAC 1108
 QY 380 rGlyIleCysArgCysProCysTrpLeuTrpMetArgCysThrArgThrCysLeuGlyTr 400
 DB 1109 AGAATTTGACAGTCTCTCTTACGCTGTGAGAAATGACATGACACTTACCAAGATG 1168
 QY 400 pSerArgThrCysThrGly-----IleSerProSerHisSerProThrGlnLeuLeu 418
 DB 1169 AAGAGGACATTTACAGGAGAGATGCTGCTGCGCCACATACATTTGAGAGTGTCT 1228
 QY 418 uProProSerSerValGluProProLeuTrpAlaProProSer 432
 DB 1229 G-----ACTACAGCATACCTCATTTCTGGGCTCAACAGCTCC 1265
 RESULT 13
 AK096459
 LOCUS 2500 bp mRNA linear PRI 15-JUL-2002
 DEFINITION Homo sapiens cDNA FLJ39140 fls, clone NTONG2009233, moderately
 similar to BESTROPHIN.
 ACCESSION AK096459

VERSION AK096459.1 GI:21755962
KEYWORDS Oligo cloning; full insert sequence;
SOURCE Homo sapiens tongue cDNA to mRNA, clone: NTONG2009233.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS

1 Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Nagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahori, K., Masuno, Y., Nagai, K. and Isogai, T.
Unpublished
NEDO human cDNA sequencing project
2 (bases 1 to 2500)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - 6' - 3' - end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology, RAB, Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

COMMENT

FEATURES

source

1. 2500
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NTONG2009233"
/tissue_type="tongue"
/clone_id="NTONG2"
/note="cloning vector: pME18SFL3"
229..2235
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04797.1"
/db_xref="GI:21755963"
/translation="MTVTYSKVANATFEFGFRLILKMGSIYKLLYEETVEAVLYT
AISTYRLILGVQKRYFEKLSICDRAEQIPVTEVGVVTVLVNMMQFNVLPH
PDRMLPFISSVHSGDEHGLRLRLIMAYVLTSLIRSVSTAYKRFPTMDHVEA
GFMTDERKLEFNLHSPHLKYVPIFGNATKARNGRIRDSVDLOSMTENRYA
SMCSLIFGDMVGILVYTOYVLAITYFPACILGRFDPITGAGADLDLYPIR
TLOEFFYAGMLKVAIBOLINFGEDDDFEFNWCIDRLIOVSLAVEMHDLIPID
DIYWDSDAARPYTLAADVCIPFSLGTSYVOMSGSDFPDEWLMYERKGGHSMI
RKYKRFISAHEHPSRRRSYRQTSDSMELPRDSDSPADLDIVSRNPRASPTW
KSCFEPGSPTLHFSMELSTIRETSQTSLSLTPSSVTSIKILYPIVETITTA
EAVPVTSGGYHDSATSIISSEFGVOPSTEEQQCGMSILSSEKETPPGGSPPT
VSAEAENINFCNDEDPDGLFKRSLDGLGSSSTSLGNLSPDMSQAPALLIDTETS
SEISGIVIVAGSVSSDMKLYMELDKETDILNELKETEESPK"

BASE COUNT 689 a 632 c 563 g 616 t

ALIGNMENT SCORES:

Pred. No.: 9, 43e-58 Length: 2500
Score: 1063.00 Matches: 217
Percent Similarity: 60.00% Conservative: 44
Best Local Similarity: 49.89% Mismatches: 101
Query Match: 45.23% Indels: 73
DB: 9 Gaps: 3

US-09-622-964-5 (1-435) x AK096459 (1-2500)

QY 1 MetThrIleFhTyrThSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 229 ATGACTGCATCTACTCCAGTAAGTAGCAATCTTTTTCGATTTCATAGCTTA 288

QY 21 LeuLeuCysTrpArgGlySerIleTyrLysLeuLeuTyrGlyGluPheIlePheLeu 40
DB 289 CTCCTCAAGTGGAGAGGAGCAGCATCTCAAACTACTGTACAGGAAATTATGTTTGGT 348
QY 41 LeuCysTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGluGlnGlnLeu 60
DB 349 GTCTCTTTATACGCAATTAAGTTGGTGTAACAATGTGTACTTACAGAGACTCCAAACGT 408
QY 61 MetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
DB 409 TACTTGAATAATATTCATATTACTGTGACAGATATGCTGAACAAATTCACACTT 468
QY 81 ValLeuGlyPheTyrValThrLeuValAlaThrArgTyrPhePheGlnTyrLysAsnLeu 100
DB 469 GTGCTGGGTTTATGTTACTCTGCTGTAACCATGTGTGACCATGTTGTGAAATTG 528
QY 101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGlyLysAspGluGln 120
DB 529 CCTTGCCAGACAGGCTTAATGTTCTCATCTCTAGCAAGTGTTCACGAAAGGAGAGCAG 588
QY 121 GlyArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeu 140
DB 589 GGGGCGCTGTAGAAAGACGCTGATGCTGCTCAATCTCACCTCCCTCATCTTT 648
QY 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
DB 649 CGCTGGTGAGCACTGCTGTGACAAAGATTTCCACANTGACACCGAGTGTGAAGCA 708
QY 161 GlyPheMetThrProAlaGluHisLysGlnLeuGlyLysLeuSerLeuProHisAsnMet 180
DB 709 GGTTTATGACACACAGATGAAGAAATTTATCAACCCCAAGTCTCCTCATCTGAA 768
QY 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
DB 769 TATTTGGTTCACATCTATCTGTTTGAAATCTTGCACTGAACCCGGAATGAAGTGA 828
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
DB 829 ATCAGAGACAGTGTATGATCTGCAATCATGATGATGATGATGATGATGATGATGAT 888
QY 221 CysGlyHisLeuLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnVal 240
DB 889 TGCACGCTCTTATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948
QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB 949 ACTCTTGCTGCTATACCTCTCTCTTGGCGTCTGATGAGCCAGCTTTTGGATGCC 1008
QY 261 AlaLysAlaTyrProGlyHisGlyLeuAspLeuValAlaProValPheThrPheLeuGln 280
DB 1009 ACCAAAGGCTACGAGGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
QY 281 PhePhePheTyrValGlyTyrLeuLysValGlyLeuSerArgAlaLeuLeuGlyTyrArg 300
DB 1069 TTTCTTCTTATGACGAGGTGGCTTAAGGTA 1098
QY 301 HisGlyIleArgGlyHisGlyGlnGlnLeuLeuGluThrArgMetGlnCysGlnGluArg 320
DB 1098 1098
QY 321 LysValSerArgValGluSerSerGlnAlaTrpTrpArgThrProValIleProAlaThr 340
DB 1098 1098
QY 341 ArgGluAlaGluAlaGlyLysSerLeuGluProGlyArgArgArgLeuTrpTrpGln-Se 360
DB 1099 1099
QY 360 rSerSerSerThrProLeuGluArgMetMetIleLeuArgProThrGlyLeuSerThr 380
DB 1105 CAGCTTATCAACCCCTTTGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1164

QY 380 rgllylecyasargcysprocystrpleutrpmetargcysthargthrcysleuglytr 400
 Db 1165 AGAATTGAGGCTCTCTTTAGCTGTGAGCAAAATGACATGACCTTACCCAGATG 1224
 QY 400 pseratgthrcysrthgyl-----lleserproserhiserprothrcinleu 418
 Db 1225 AAGAGGACATTTACTGCGAGAGATTCGTGCTGCGCCACATACATTTGGAGCGTCT 1284
 QY 418 uproserSerValglupProleutrpAlaProProSer 432
 Db 1285 G-----ACTACTGACATACCTCATTTCTGCGGTCAACAGCTCC 1321
 RESULT 14
 BC031186 1956 bp mRNA 1linear ROD 07-AUG-2002
 LOCUS Mus musculus, clone MGC:37621 IMAGE:4989959, mRNA, complete cds.
 DEFINITION BC031186
 ACCESSION BC031186.1 GI:21411099
 VERSION
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1956)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
 USA
 NIH-MGC Project URL: http://mgc.ncl.nih.gov
 REMARK Contact: MGC help desk
 COMMENT Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www.shgc.stanford.edu
 Contact: (Dickson, Mark) mcdickaxl.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Series: IRM Plate: 58 Row: 9 Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source Location/Qualifiers
 1..1956
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:37621 IMAGE:4989959"
 /tissue_type="Colon, normal, 5 month old male mouse."
 /clone_1lb="NCT-CCAP-Co24"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 176..1573
 /codon_start=1
 /product="Unknown (protein for MGC:37621)."
 /protein_id="AAH31186.1"
 /db_xref="GI:21411100"
 /db_xref="locusID:212989"
 /translation="MALSAAYRFLLAERQKRYFKLVICDOYASLIPVSEVLAGVYV
 LVHRMNOYICMLPDAIMCIVAGTVGRDGRGLRYRTLMKRYAGLSAVILTSYST
 AVFKRPTIDHVEAGFVTRERKKFEMUNSYKVMPCWVSSIAOARREGIR
 NSAKLLELVNVRKCGMLFHYDMISIPLYTYVITAYISYFLACILIGQFLDPA
 OGKXKHTDLCYPTITLQFFRYAGLVAEQLINPREDDDDEFENFLIDNKEQVSM
 LAVDMTDLAEKDLTWDAAEAPAYTAATFALLOOPSFGSGSTFDLIDAKEDQFO
 RUDGVDPLGAEHGDIFQRLPAGASVPTAATLRLKRSKCSVSEASTAASCAG
 AADGCGVCGCGDPLDPSLRPELEPACPEPAPIDGPPEPTTVSISGPNAPAP

BASE COUNT 411 a 577 c 530 g 438 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,02e-57 Length: 1956
 Score: 1060.50 Matches: 219
 Percent Similarity: 60.60% Conserved: 44
 Best Local Similarity: 50.46% Mismatches: 99
 Query Match: 45.13% Indels: 72
 DB: 10 Gaps: 4
 US-09-622-964-5 (1-435) x BC031186 (1-1956)
 QY 1 MetThrIleThrThyThrSerGlnValAlaAsnAlaArgLeuIleSerPheSerArgLeu 20
 Db 47 ATGACCGTCACCTACACAGCAGCAGAGGCGAATGCGGCTTCGGTGGCTTCGACAGCTG 106
 QY 21 LeuLeuCystrPrArgIleSerIleTyrlsLeuLeuTyrlGlyLubLeuIlePheLeu 40
 Db 107 CTGCTCTGTGGCGGAGAGCATCTACAGCTCTGTGGAGAGCTGTATGTTCTTC 166
 QY 41 LeuCystrTyrlIleIleArgPheIleTyrlArgLeuAlaLeuThrgLugInGlnLeu 60
 Db 167 GGACTCTACATGACACTAAGCGCGCCCTATCGCTTACTGTGCAGAAAGCAGAGCGC 226
 QY 61 MetPheGlnLysLeuThrLeuTyrlCysAspSerTyrlIleGlnLeuIleProIleSerPhe 80
 Db 227 TACTTGGAGAGGTTGTCTATATACGACAGCAGTACGACAGCTCATCCCGCTCTTTC 286
 QY 81 ValLeuGlyPheTyrlValThrLeuValAlaThrArgTrpPheAsnGlnTyrlGlnLeu 100
 Db 287 GACTTGGCTTCTAGCTGACTGTGGTGCATCGCTGTGGAGACAGTACCTATCCATG 346
 QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlnLysAspGln 120
 Db 347 CCTTCCCGACGACATCTATGTCATATGCTGCGTGCAGCCGTGATGCGAGCATCA 406
 QY 121 GlyArgLeuLeuArgThrLeuIleArgTyrlAlaAsnLeuGlnLysValLeuIleLeu 140
 Db 407 GGCCGCTTACCGCGCGCGCTCATGCCCTACGAGGCTCTCCCGGTGCTGATCTT 466
 QY 141 ArgSerValSerThrAlaValTyrlsArgPheProSerAlaGlnHisLeuValGlnAla 160
 Db 467 CATTCTGTAGACAGACAGCTTCAACAGCTTCCACATATAGACAGCTGTGCGAGGT 526
 QY 161 GlyPheMetThrProAlaGlnHisLysGlnLeuGlnLysLeuSerLeuProHisAsnMet 180
 Db 527 GATTATATGACCCGAGAAAGCGCAAGGTTCCAGAACTTGAATTCCTCAACAA 586
 QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
 Db 587 TACTGGGTCCCTCGGTATGTTTCGAGCTTCGACCGCAGCGACGCGGAGAGGGGCC 646
 QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnTrpLeuArgTrgIn 220
 Db 647 ATCCCGGACACAGCGCCCTAAGTTGCTACAGAGAGACTGAATGTGTTGGAGCA 706
 QY 221 CysGlnHisLeuTyrlAlaTyrlAspTrpIleSerIleProLeuValTyrlThrcInVal 240
 Db 707 TGTGGAGTGTCTTCACTACAGCTGATACCAATCCCTGTCTACACCCAGGTAGTC 766
 QY 241 ThrValAlaValTyrlSerPhePheLeuThrcysLeuValGlyArgGlnPheLeuAsnPro 260
 Db 767 ACTATCGCAGGTATACAGTACTTCTGTGCTGCTCCTCAGCGCGAGTCTCTACACCT 826
 QY 261 AlaLysAlaTyrlProGlyHisLysLeuAsnLeuValAlaProValPheThrcPheLeuGln 280
 Db 827 GCACAGGCTCAACAAAGACACACCTGACCTATGCGATGCCATCTTACCTTGTGCGAG 886
 QY 281 PhePhePheTyrlValGlyTrpLeuTyrlValGlyLeuSerArgAlaLeuLeuGlyTrpArg 300
 Db 887 TTCTTCTTCTACGCGCTGCTTACGTA----- 916

[illegible]

```
QY 201 ILeaRgAspProIleuLeuInSerLeuLeuAnGluMetAsnThrLeuArgThrGln 220
Db 660 ATCCGACACAAACAGTGCCTAAAGTCTACTAGAGAGCGAATGTGTTCGAGACAA 719
QY 221 CysGlyHisLeuTyraIaTyraSPTrpIleSerIleProLeuValTyrThrGlnVal 240
Db 720 TGTGGATGCTGTTCTACGACTGATGATACCCCTCGTACACCCAGTAGTC 779
QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 780 ACTATCCAGTGTACAGCTACTCTTGTGCTTCCTCATCGCCGCTGCTTCCAGACCT 839
QY 261 AlaLysAlaTyrProGlyHisGluLeuAspLeuValProValPheThrPheLeuGln 280
Db 840 GCACAGGGCTACAAGACACACACCTGACCTATGCCATCCATCTCACCCTGCTCAG 899
QY 281 PhePhePheTyrValGlyTrpLeuLysValGlyLeuSerArgAlaLeuLeuGlyTrpArg 300
Db 900 TTTCTTCTTACGCTGCTGCTTAAGTA----- 929
QY 301 HisGlyGlnArgGlyHisGlyGlnGluLeuGluThrArgMetGlnCysGlnGluArg 320
Db 930 -----GCAGACACCTTATT----- 944
QY 321 LysValSerArgValGluSerSerGlnAlaTrpTrpArgThrProValIleProAlaThr 340
Db 944 ----- 944
QY 341 ArgGluAlaGluAlaGlyGluSerLeuGluProGlyArgArgLeuTrpTrpGlnSer 360
Db 944 ----- 944
QY 361 SerSerSerThrPro-LeuGluArgMetMetIleLeuArgProThrGlyLeuSerTh 380
Db 945 -----AACCTTTTGGAGAGAGACGACGACGACTTGTAGACCAACTTCTTATTGAC 995
QY 380 rGlyIleCysArgCysProCysTrpLeuTrpMetArgCysThrArgThrCysLeuGlyTr 400
Db 996 CGCACTTCCAGGTGTCATGCTAGCTGTAGATGATGATGACGACTTGGCCATGCTG 1055
QY 400 pSerArgThrCysThrGlyIleSerProSerHisSerProThrGlnLeuLeuProP 420
Db 1056 GAGAAAGATCTATCTGAGTGCAGACAGAAAGCTCGCGCCCTTACACCG---CGGCCACC 1112
QY 420 oSerSerVal-----GluProProLeuTrpAlaProPro 431
Db 1113 GCTTTCCTGCTTACAGACGCCCTCTCCAGGGGCTCCACCT 1152
```

Search completed: July 27, 2003, 15:00:47
Job time : 3689.65 secs